The Biometric Society

BIOMETRICS

FOUNDED BY THE BIOMETRICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION

TABLE OF CONTENTS

Number 4 **December 1957** **Volume 13**

THE BIOMETRIC SOCIETY

General Officers

President, E. A. Cornish; *Secretary*, M. J. R. Healy; *Treasurer*, A. W. Kimball; *Council*, F. J. Anscombe, Claudio Barigozzi, L. L. Cavalli-Sforza, W. G. Cochran, G. M. Cox, Georges Darmois, B. B. Day, D. J. Finney, J. H. Gaddum, M. P. Geppert, Americo Groszmann, P. C. Mahalanobis, Donald Mainland, Leonard Martin, Motosaburo Masuyama, P. A. P. Moran, Jerzy Neyman, C. R. Rao, E. J. Williams, Frank Yates, W. J. Youden.

Regional Officers

Eastern North American Region: *Regional President*, B. Harshbarger; *Secretary-Treasurer*, A. M. Dutton. British Region: *Regional President*, D. J. Finney; *Secretary*, C. C. Spicer; *Treasurer*, A. R. G. Owen. Western North American Region: *Regional President*, D. G. Chapman; *Secretary-Treasurer*, Elizabeth Vaughan. Australasian Region: *Regional President*, E. J. Williams; *Secretary*, G. S. Watson; *Treasurer*, G. A. McIntyre. French Region: *Regional President*, Eugene Morice; *Secretary-Treasurer*, Daniel Schwartz. Belgian Region: *Regional President*, R. Laurent; *Secretary*, Leopold Martin; *Treasurer*, A. H. L. Rotti. Italian Region: *Regional President*, Gustavo Barbensi; *Secretary*, L. L. Cavalli-Sforza; *Treasurer*, R. E. Scossiroli. German Region: *Regional President*, O. Heinisch; *Secretary-Treasurer*, Wilhelm Ludwig. Brazilian Region: *Regional President*, C. G. Fraga, Jr.; *Secretary*, P. Mello Freire; *Treasurer*, A. Groszmann.

National Secretaries

Denmark, N. F. Gjeddebaek; India, K. Kishen; Japan, M. Hatamura; The Netherlands, E. van der Laan; Sweden, H. A. O. Wold; Switzerland, H. L. LeRoy.

Editorial Board
Biometrics

Editor: Ralph A. Bradley; *Editorial Associates and Committee Members:* C. I. Bliss, Irwin Bross, E. A. Cornish, S. Lee Crump, W. J. Dixon, Mary Elveback, J. W. Hopkins, O. Kempthorne, Leopold Martin, Horace W. Norton, G. W. Snedecor and Georges Teissier. *Managing Editor:* Ralph A. Bradley.

The Biometric Society is an international society devoted to the mathematical and statistical aspects of biology. Biologists, mathematicians, statisticians and others interested in its objectives are invited to become members. Through its regional organizations the Society sponsors regional and local meetings. National secretaries serve the interests of members in Denmark, India, Japan, the Netherlands, Sweden and Switzerland, and there are many members at large.

Rates (in U.S.A. currency) for full membership in the Society for 1957 including dues and a subscription to this journal are: for residents of Canada and the United States $7.00 (Student Membership $4.00); for others $4.50 (Student Membership $2.75). Members of the American Statistical Association who are currently subscribing to Biometrics through that organization may become members of the Biometric Society on payment of $3.00 annual dues if resident in the United States or Canada, and of $1.75 annual dues if resident elsewhere. Information concerning the Society can be obtained from its Secretary, M. J. R. Healy, Statistics Department, Rothamsted Experimental Station, Harpenden, Herts, England.

Annual subscription for non-members of the Biometric Society is $7.00, payable to the Managing Editor, *Biometrics*, Dept. of Statistics, Virginia Polytechnic Institute, Blacksburg, Virgina.

Second-class mailing privileges authorized at Knoxville, Tenn. Additional entry at Richmond, Va. Business Office: 509 West Hill Road, Knoxville 19, Tennessee. *Biometrics* is published quarterly—in March, June, September and December.

A MODEL FOR THE DETERIORATION IN STRENGTH OF MATERIALS DUE TO FUNGAL ATTACK

J. S. Maritz

Transvaal and Orange Free State Chamber of Mines Research Laboratories, Johannesburg, South Africa

INTRODUCTION

Many materials used by industries are subject to fungal attack and the conditions to which they are subjected are frequently favourable to the growth of fungi. The result of fungal attack is usually that the material gradually loses its strength and eventually disintegrates; but it may actually become useless long before such an advanced stage is reached. On the other hand, due to slow decrease in strength the material may serve its purpose for a long time even though it is being attacked by fungi. Laboratory studies of the rate of decrease in strength under different conditions are, therefore, of considerable importance.

Development of a mathematical model for the loss of strength due to fungal attack seems desirable for at least two reasons. They are:

(i) the results of the laboratory experiments are usually observations on a number of random variables and a mathematical model provides a description and summary of a set of data which can be very useful, particularly when different sets are to be compared with each other.

(ii) as many of these investigations are rather lengthy, a sound model can be of great use in the efficient planning of experiments.

The basic assumptions on which the model in this paper rests are: if a specimen of material is exposed (to fungal attack), a certain period of time will elapse before the attack actually starts and, once it starts, the strength decreases gradually to zero. Both the time of 'survival' and the rate at which the strength decreases may be random variables.

A specific example, in which connection this approach has been successful, is encountered on the South African gold and coal mines where various fabrics are used underground for dust filtering and other purposes. An example of a laboratory experiment with Cotton Duck

in which tensile strengths of pieces of standard shape and size were measured is given in a later section.

MATHEMATICAL MODEL: GENERAL FORM

Suppose that a number of sample pieces from the same batch of a material are exposed to fungal attack under conditions as nearly identical as possible. The attack on individual samples will actually start after time intervals t_1 , t_2 , $\cdots$, where $t = 0$ corresponds to the start of the experiment. Each t_i may be regarded as a realization of a random variable t with probability density $g(t)dt$, so that the expected proportion of samples remaining with no attack after time t is

$$1 - \int_0^t g(t)\, dt = 1 - G(t).$$

The initial (i.e. unexposed) strength will vary from one sample to another, but, for the moment, we consider only samples with the same initial strength x_0 . Take one on which attack starts at time t_1 , i.e. during the interval $t_1 \pm \frac{1}{2}\delta t_1$. At t_2 ($t_2 > t_1$), due to the attack the strength has decreased to $x_0 f(t_2 - t_1)$. This function $f(t_2 - t_1)$ must satisfy two conditions:

(a) $f(t_2 - t_1) = 1$ for $t_2 = t_1$

(b) $f(t_2 - t_1) \to 0$ as t_2 increases.

The form of f may itself be a random variable, but we make the simplifying assumption that it remains fixed and that its value for fixed $(t_2 - t_1)$ is a random variable only by virtue of its dependence on a constant a which is a random variable. This means that the samples on which attack starts at t_1 will at t_2 have strengths $x_0 f(t_2 - t_1 , a_1)$, $x_0 f(t_2 - t_1 , a_2)$ $\cdots$ with expected mean value

$$E_a[x_0 f(t_2 - t_1 , a)] = x_0 E_a[f(t_2 - t_1 , a)].$$

We are now able to calculate the expected mean value at time t of all samples with initial strength x_0 on the assumption that the random variables x and t are independent: the total proportion of attacked samples after time t is $G(t)$. The proportion of these attacked during the interval $\tau \pm \frac{1}{2}\delta\tau$ is $g(\tau)\delta\tau/G(t)$, and the mean strength of these at time $t(> \tau)$ is $x_0 E_a\,[f(t - \tau, a)]$. Therefore the mean of all attacked samples at time t is

$$[x_0/G(t)] \int_0^t E_a[f(t - \tau), a]g(\tau)\, d\tau,$$

and it follows that the overall mean at t is

$$x_0 \int_0^t E_a[f(t - \tau, a)]g(\tau)\, d\tau + x_0[1 - G(t)]. \tag{1}$$

Assuming that the random variables a and x are independent, we find that the expected mean strength at time t is

$$E(y_t) = E(x) \int_0^t E_a[f(t - \tau, a)]g(\tau)\, d\tau + E(x)[1 - G(t)], \tag{2}$$

where $E(x)$ is the expected mean strength of unexposed samples.

MATHEMATICAL MODEL: SPECIAL FORMS

The special forms which the functions $g(t)$ and $E_a[f(t_2 - t_1, a)]$ may assume are, of course, numerous and we will give only two special forms of $E(y_t)$ resulting from what appear to be reasonable assumptions about the two functions.

The function $g(t)$. The distribution of t must be continuous starting at $t = 0$ and covering the range $0 \to \infty$. A simple and plausible choice for $g(t)$ is

$$g(t) = \beta e^{-\beta t}. \tag{3}$$

This distribution has only one parameter, β, with the 'physical' meaning that the mean time of survival before attack starts is $1/\beta$. We were influenced in this choice of $g(t)$ by the knowledge that this type of distribution has been found to fit a large variety of 'life' distributions, e.g., life of electron tubes [Epstein 1954].

The function $E_a[f(t_2 - t_1, a)]$. In considering this function we must really take account of two functions, i.e. f and the probability density $p(a)$ of the random variable a. We have examined various forms of f and found

$$f(t_2 - t_1, a) = e^{-a(t_2 - t_1)} \tag{4}$$

one of the most easily handled and also practically successful.

The probability distribution of a. Unless fairly simple forms for $p(a)$ are used in conjunction with (4), the resulting expression for $E_a[f(t_2 - t_1, a)]$ can be very awkward indeed. Following are two examples of special forms of $E(y_t)$ obtained by using (3), (4), and two very simple forms of $p(a)$:

Example (i)

We note that, substituting (3) and (4) in (2) gives the result

$$E(y_t) = E(x) \int_0^\infty \beta \frac{(e^{-\beta t} - e^{-at})}{(a - \beta)} p(a)\, da + E(x)e^{-\beta t}. \tag{5}$$

If we put

$$p(a) = \delta(a - \alpha), \tag{6}$$

the Dirac delta function, i.e. a assumes the value α with probability 1, (5) becomes

$$E(y_t) = \frac{E(x)}{(\alpha - \beta)} [\alpha e^{-\beta t} - \beta e^{-\alpha t}]. \tag{7}$$

This is a simple expression which may readily be 'fitted' to observed data, and we have found it to be applicable to many of our laboratory experimental results. An example will be given in a later section.

*Example (ii)**

If we assume, instead of (6), that the distribution of a is located at α and has a small variance σ^2, expression (2) may be written approximately as

$$E(y_t) \doteqdot E(x) \int_0^t \left[f(t - \tau, \alpha) + \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial \alpha^2}(t - \tau, \alpha) \right] g(\tau)\, d\tau + E(x)[1 - G(t)], \tag{8}$$

and substituting (3) and (4) we have

$$E(y_t) \doteqdot \frac{E(x)}{(\alpha - \beta)} [\alpha e^{-\beta t} - \beta e^{-\alpha t}] + \frac{\sigma^2 \beta E(x)}{(\alpha - \beta)} \left[\frac{1}{(\alpha - \beta)^2} (e^{-\beta t} - e^{-\alpha t}) - \tfrac{1}{2} e^{-\alpha t} \left(t^2 + \frac{2t}{(\alpha - \beta)} \right) \right]. \tag{9}$$

This is a more flexible expression than (7), as can be seen more clearly if we consider the special case $\beta = \alpha$ which gives for (7) and (9) respectively,

$$E(y_t) = E(x) e^{-\alpha t}(1 + \alpha t)$$

and

$$E(y_t) = E(x) e^{-\alpha t} \left(1 + \alpha t + \frac{\alpha \sigma^2}{6} t^3 \right).$$

As the form (7) appears, despite its simplicity, to be of real practical use, the following section gives some detail about estimating its param eters.

---

*I would like to acknowledge my indebtedness to one of the referees for drawing my attention to this example.

METHODS OF ESTIMATING THE PARAMETERS IN EQUATION (7)

The parameters are $E(x) = \xi_1$, α and β. As our main concern so far has been to see whether the model fits our observations at all, we have not gone into the question of estimation very carefully and can suggest only some crude methods of estimating the parameters.

Considering the way in which the model is built up, it is clear that in some circumstances ξ_1 and β may be estimated without using the observed regression of y_t on t.

Estimate of ξ_1 . By definition ξ_1 is the expected mean value of unexposed unattacked samples so that if we have N such samples with strengths $x_1 \cdots x_N$ an estimate of ξ_1 is simply

$$\hat{\xi}_1 = \sum x_i/N. \tag{10}$$

As it is usually true that N can be made large without much trouble or expense this procedure may be recommended.

Estimate of β. Sometimes it is possible (e.g. by microscopic examination) to discriminate between attacked and unattacked samples before they are, as is often the case, tested to destruction. When this can be done, a series p_1 , p_2 , $\cdots$ of observed proportions of attacked samples at times t_1 , t_2 $\cdots$ is found. According to (3)

$$E(p_i) = 1 - e^{-\beta t_i}. \tag{11}$$

Fitting the expression (11) through the observed p_i by, for example, the maximum likelihood method gives an estimate of β.

Estimate of α *when* β *and* ξ_1 *are known.* A weighted least squares procedure appears to be advisable. If the observed values of y_t at t_i (> 0) are $y_{i1} \cdots y_{in_i}$ and the weights are $w_1 \cdots w_k$, we minimize

$$S_0 = \sum w_i \left[y_{i.} - \frac{\xi_1}{(\alpha - \beta)} (\alpha e^{-\beta t_i} - \beta e^{-\alpha t_i}) \right]^2 \tag{12}$$

with respect to α, where $y_{i.} = \sum_i y_{ij}/n_i$, and the result is

$$\sum w_i [y_{i.} - \xi_1 Y_i][Y_i - (e^{-\beta t_i} + \beta t_i e^{-\alpha t_i})] = 0 \tag{13}$$

where $\xi_1 Y_i = E(y_{t_i})$.

If ξ_1 and β are unknown but estimated as above, those estimates may be substituted in (13), which procedure will lead to a consistent estimate of α.

Joint estimation of α *and* β *by weighted least squares.* When a separate estimate of β cannot be obtained, α and β may be estimated jointly by minimizing S_0 [see equation (12)] with regard to both α and β.

One difficulty is, however, that α and β are interchangeable in the expression (7). It may be resolved by making use of the fact that

$$E(y_i^2) = \frac{E(x^2)}{(2\alpha - \beta)} [2\alpha e^{-\beta t} - \beta e^{-2\alpha t}] = \xi_2 \Phi(2\alpha, \beta, t), \tag{14}$$

where $\xi_2 = E(x^2)$, which may be derived in the same way as the value of $E(y_i)$. For, it may be shown that a consistent estimate of α is obtained if, of the two possible estimates of α, that one is chosen which gives the smaller difference, in some sense between $\xi_2 \Phi(2\hat{\alpha}, \hat{\beta})$ and the observed values $\sum y_{ij}^2/n_i$. Even a crude rule such as: choose $\alpha = \hat{\alpha}$ or $\hat{\beta}$ according to whether $u < v$ or $u > v$ where

$$\begin{aligned} u &= \sum_i [\sum_j y_{ij}^2/n_i - \hat{\xi}_2 \Phi(2\hat{\alpha}, \hat{\beta}, t_i)] \\ v &= \sum_i [\sum_j y_{ij}^2/n_i - \hat{\xi}_2 \Phi(2\hat{\beta}, \hat{\alpha}, t_i)] \end{aligned} \tag{15}$$

can be satisfactory.

Standard errors of weighted least squares estimates of α and β. By argument similar to that used, for example, by Kendall [Kendall, 1948, p. 208] to calculate the standard errors of functions of moments, one can show that the variances and covariances of least squares estimates $\hat{\alpha}$ and $\hat{\beta}$ of α and β are approximately given by

$$\text{cov}\,(\hat{\alpha}, \hat{\beta}) = (A^{-1})B(A^{-1})', \tag{16}$$

where

$$A = \begin{bmatrix} \sum w_i \left(\frac{\partial \psi_i}{\partial \alpha}\right)^2 & \sum w_i \left(\frac{\partial \psi_i}{\partial \alpha}\right)\left(\frac{\partial \psi_i}{\partial \beta}\right) \\ \sum w_i \left(\frac{\partial \psi_i}{\partial \alpha}\right)\left(\frac{\partial \psi_i}{\partial \beta}\right) & \sum w_i \left(\frac{\partial \psi_i}{\partial \beta}\right)^2 \end{bmatrix},$$

$$B = \begin{bmatrix} \sum \left(\frac{\partial \psi_i}{\partial \alpha}\right)^2 w_i^2 \sigma_i^2 & \sum \left(\frac{\partial \psi_i}{\partial \alpha}\right)\left(\frac{\partial \psi_i}{\partial \beta}\right) w_i^2 \sigma_i^2 \\ \sum \left(\frac{\partial \psi_i}{\partial \alpha}\right)\left(\frac{\partial \psi_i}{\partial \beta}\right) w_i^2 \sigma_i^2 & \sum \left(\frac{\partial \psi_i}{\partial \beta}\right)^2 w_i^2 \sigma_i^2 \end{bmatrix},$$

$$\psi_i = \frac{\xi_1}{(\alpha - \beta)} (\alpha e^{-\beta t_i} - \beta e^{-\alpha t_i}) = \xi_1 \Phi(\alpha, \beta, t_i),$$

and $\sigma_i^2 = \text{var}\,(y_i.)$.

If w_i is chosen equal to $1/\sigma_i^2$, the result is simply

$$\text{cov}\,(\hat{\alpha}, \hat{\beta}) = A^{-1}. \tag{17}$$

To take account of the variance of the estimate $\hat{\xi}_1$ of ξ_1, the elements of B must be increased by amounts

$$\left(\sum w_i \Phi \cdot \frac{\partial \psi_i}{\partial \alpha}\right)^2 \text{var } (\hat{\xi}_1), \text{ etc.}$$

Finally, the variance of $\hat{E}(y_t)$ is approximately (for ξ_1 known) given by

$$\text{var } [\hat{E}(y_{t_i})] = \left(\frac{\partial \psi_i}{\partial \alpha}, \frac{\partial \psi_i}{\partial \beta}\right) A^{-1} \left(\frac{\partial \psi_i}{\partial \alpha}, \frac{\partial \psi_i}{\partial \beta}\right)'. \tag{18}$$

PRACTICAL EXAMPLE

Standard size pieces of Cotton Duck treated with copper naphthenate were exposed to attack by a test fungus (T.R.L. 255 Chaetomium Globosum). Altogether 30 pieces were exposed and by random choice they were tested for tensile strength in groups of six each after 13, 26, 39, 52, and 65 days of exposure. The tensile strengths of ten unexposed pieces were measured. The individual tensile strength values (in arbitrary units) are given in Table 1.

TABLE 1

TENSILE STRENGTHS OF COTTON DUCK

	Period of exposure in arbitrary time units				
Unexposed	1	2	3	4	5
100.5	76.0	24.2	13.0	1.5	0.0
88.2	67.3	33.3	3.5	0.0	0.0
107.5	66.5	28.3	13.0	4.8	0.0
102.5	58.6	37.3	8.3	2.5	2.5
96.0	67.3	34.5	5.0	0.5	0.0
107.8	74.8	22.0	5.0	5.8	2.0
94.0					
103.5					
98.9					
99.0					

All the pieces examined after 13 days of exposure were attacked so that these data were not considered suitable for estimating β as suggested by equation (11). As estimate of ξ_1, the mean of the unexposed pieces was used, i.e. $\hat{\xi}_1 = 99.79$. Estimates of α and β were then found by minimizing S_0 with respect to α and β, where $\hat{\xi}_1$ was substituted for ξ_1 and the weights w_i were taken as $w_i = n_i/s_i^2$ where

s_i^2 is the observed variance of the strengths at time t_i . The result was that the two values of $\hat{\alpha}$ and $\hat{\beta}$ were 1.30 and 1.15. The observations and the graph $\hat{\xi}_1\Phi(1.30, 1.15)$ are plotted in the accompanying figure and it is clearly a very good fit. A plot of the squares of the observations and the curves $\hat{\xi}_2\Phi(2.60, 1.15)$ and $\hat{\xi}_2\Phi(1.30, 2.30)$, showed by inspection that we should take $\hat{\alpha} = 1.30$, $\hat{\beta} = 1.15$.

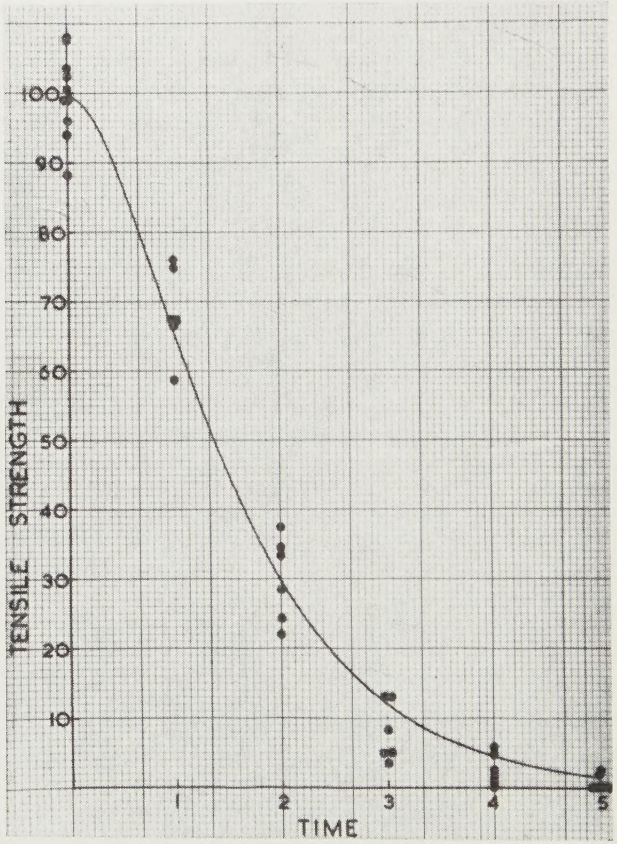


FIGURE 1

TENSILE STRENGTHS OF COTTON DUCK

The variances and covariances of $\hat{\alpha}$ and $\hat{\beta}$ calculated as suggested above turned out to be

$$\text{cov}\,(\hat{\alpha}, \hat{\beta}) \approx \begin{bmatrix} 1.323 & -1.003 \\ -1.003 & .763 \end{bmatrix}.$$

These values are surprisingly large, due, we believe, to the fact that in this example $\hat{\alpha}$ is very nearly equal to $\hat{\beta}$. However, due to the high negative correlation between $\hat{\alpha}$ and $\hat{\beta}$, using equation (18) shows that $E(y_{t_1})$ may be estimated with reasonable precision. For example at $t = 1$, s.d. $\hat{E}(y_t) = 2.12$. Of course, bringing into the calculation the variance of $\hat{\xi}_1$ increases this standard deviation and the result is s.d. $\hat{E}(y_t) = 3.18$. At $t = 5$ the two standard deviations are respectively .28 and .32. No details are given of the calculation incorporating var $(\hat{\xi}_1)$ because normally it is easy to make N so large that var $(\hat{\xi}_1)$ may be made negligible.

ACKNOWLEDGEMENTS

I wish to thank the Transvaal and Orange Free State Chamber of Mines for permission to publish this paper, and the staff of the Timber Research Laboratory who presented me with the problem and carried out the experiment reported above.

REFERENCES

(1) Epstein, B. [1954]. Truncated life tests in the exponential case. *Ann. Math. Stat.*, *25*: 555–564.
(2) Kendall, M. G. [1948]. *The Advanced Theory of Statistics*. 4th ed. London: Charles Griffin and Company, Ltd.

OPTIMUM GROUP SIZE IN PROGENY TESTING AND FAMILY SELECTION

ALAN ROBERTSON
Institute of Animal Genetics, Edinburgh, Scotland

The research worker concerned with animal breeding problems is frequently asked the question "How many progeny do I need to test a sire adequately?" This is an awkward, if not impossible, question to answer because it depends on the meaning of the word "adequately." The more progeny the more accurate the assessment—the level of accuracy that is considered adequate is a matter of subjective judgment. But, by making the content of the question a little more practical and precise, it can be put in forms to which a useful solution is possible.

In any progeny-testing programme (this implying a breeding programme, relying on progeny-testing for improvement, rather than a method of evaluating sires), two factors external to the actual testing procedure will define the programme. They are the facilities available (limiting the total number of progeny which can be tested in any generation) and the number of sires that will be selected each generation. The latter will be limited by the rate of inbreeding which can be tolerated in the programme. The further variable factor in the programme is the number of sires which will be tested each generation. This, in relation to the available facilities, determines the group size. The problem then is to determine the group size which will lead to the maximum expected genetic superiority of the selected sires. This is a balance of choice against accuracy. The more sires are tested the greater will be the choice, but the less accurate the information on which the choice is made, because the number of progeny per sire is less. This problem can be solved within the framework of statistical genetics.

Within this framework, the expected genetic superiority of chosen sires is then given by

$$\Delta G = \bar{\imath} r_{IG} \sigma_g$$

where $\bar{\imath}$ is the selection intensity (in standard units) corresponding to the proportion of sires selected, r_{IG} is the correlation between the

progeny average I and the breeding value of a sire, and σ_g is the genetic standard deviation in the population. If we can assume a normal distribution of the breeding value of sires, then $\bar{i}$ can be obtained from the proportion selected from standard tables. If we know the heritability h^2 of the character on which selection is based, then if there are no non-genetic causes of differences between sire groups, we have

$$r_{IG} = \sqrt{\frac{\frac{1}{4}nh^2}{1 + \frac{1}{4}(n-1)h^2}}$$

where n is the number of progeny in the group. When we have testing facilities for N animals and a knowledge of the number of chosen sires, S, required, we can easily calculate the value of n which maximises ΔG. The results are illustrated in Table 1 for $N = 100$, $S = 2$, $h^2 = 0.25$.

TABLE 1

THE EXPECTED GENETIC SUPERIORITY OF SELECTED SIRES IN A PROGENY TESTING SCHEME

(N = total progeny measured, S = sires selected, n = offspring/sire, ΔG = genetic superiority of selected sires in genetic standard deviations).

Group size (n)	4	5	10	20
Accuracy (r_{IG})	0.46	0.50	0.63	0.76
$N = 100, S = 2$				
Sires tested	25	20	10	5
$\bar{i}$	1.74	1.64	1.27	0.83
ΔG	0.80	0.82	0.80	0.62
$N = 200, S = 4$				
Sires tested	50	40	20	10
$\bar{i}$	1.80	1.69	1.33	0.90
ΔG	0.83	0.85	0.84	0.68
$N = \infty, N/S = 50$				
Sires tested	$12.5S$	$10S$	$5S$	$2.5S$
$\bar{i}$	1.86	1.76	1.40	0.97
ΔG	0.86	0.88	0.88	0.74

ΔG passes through a maximum when n is in the region of 5. Suppose we now turn to other values of N and S, but keep the ratio of N/S constant. The values of $\bar{i}$ and ΔG are slightly altered in each case, because for a given proportion of sires selected, $\bar{i}$ increases a little as the number selected from increases. The position of the maximum is little changed. It seems then that with a considerable gain in gener-

ality, we can treat the value of N as very large and use as our new variable N/S, the total number of progeny recorded for each group that is eventually selected. This we shall call the testing ratio and give it the symbol K. It is convenient here to introduce a further variable, p, the proportion of sires selected. As the number of sires tested each generation is N/n, it follows that

$$p = \frac{S}{N/n}$$

$$= \frac{n}{K}$$

or $n = Kp$.

In terms of K and p, the problem turns out to have a general solution. To calculate $\bar{\imath}$, we assume that we are selecting from a large sample, so that

$$\bar{\imath} = \frac{z}{p}$$

where z is the ordinate of the normal curve at the point where the area cut off $= p$.
Similarly,

$$r_{IG} = \sqrt{\frac{\frac{1}{4}nh^2}{1 + \frac{1}{4}(n-1)h^2}}$$

$$= \sqrt{\frac{n}{n+a}}, \quad \left(\text{where } a = \frac{4 - h^2}{h^2}\right)$$

$$= \sqrt{\frac{p}{p + a/K}}.$$

Then

$$\Delta G = \frac{z}{p}\sqrt{\frac{p}{p + a/K}}\,\sigma_g .$$

Thus the expected superiority of the chosen sires is a function of p and K/a. For any value of K/a, we are concerned with finding the value of p which maximises ΔG. It follows that the optimum value of p will be a function only of K/a, as will the maximum value of ΔG. We have then introduced a much greater generality into the problem.

We have now to find the maximum value of ΔG or of $(z/p) \cdot \sqrt{p/(p + a/K)}$. It is algebraically much simpler to find the minimum of the square of the reciprocal of this, $[p(p + a/K)]/z^2$. In this, we

must remember that $dz/dp = x$, the abscissa of the normal curve at the point defined by p.

$$\frac{d}{dp}\left[\frac{p(p + a/K)}{z^2}\right] = \frac{1}{z^4}\left[z^2\left(2p + \frac{a}{K}\right) - p\left(p + \frac{a}{K}\right)2z\frac{dz}{dp}\right]$$

$$= \frac{1}{z^4}\left[z^2\left(2p + \frac{a}{K}\right) - 2pzx\left(p + \frac{a}{K}\right)\right]$$

$$= 0, \text{ at the minimum.}$$

This simplifies to

$$\frac{K}{a} = \frac{1}{2p}\frac{2px - z}{z - px}.$$

We have thus equated K/a to a function of p, easily calculated from tables of the normal distribution. We can thus determine the optimum value of p in terms of K/a, and therefore also the maximum value of G.

These are shown in Figures 1 and 2, in which K/a is plotted on a log scale. Figure 1 shows the optimum value of p and Figure 2 the maximum value of ΔG. Several interesting points emerge from a study of these. The first is that p increases as K/a declines but approaches asymptotically the value of 0.27 (where $2px = z$). Thus, for the optimum running of any such scheme, the intensity of selection between tested sires must be at least one in four. At values of K/a greater than 3, it is found empirically that the curve approximates closely to $p = 0.28\sqrt{a/K}$. Substituting $n = Kp$, this leads to the expression $n = 0.56\sqrt{K/h^2}$ as the optimum value for group size.

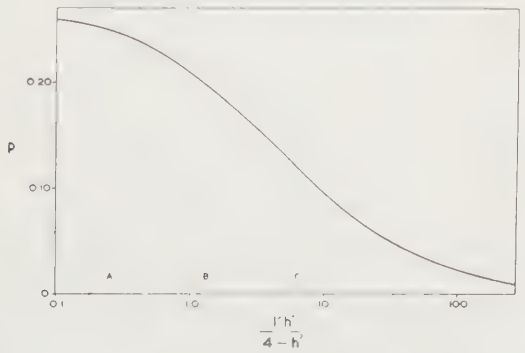


FIGURE 1

THE OPTIMUM PROPORTION OF SIRES SELECTED (p) AS A FUNCTION OF $Kh^2/(4 - h^2)$

As might be expected, ΔG_{max} increases as K/a increases. Above a value of K/a of about 5, it becomes a straight line in the plot against log

K/a. This fact, which has thus far defied algebraic check, means that a given proportionate size in K/a will produce the same arithmetic increase in the superiority of chosen sires.

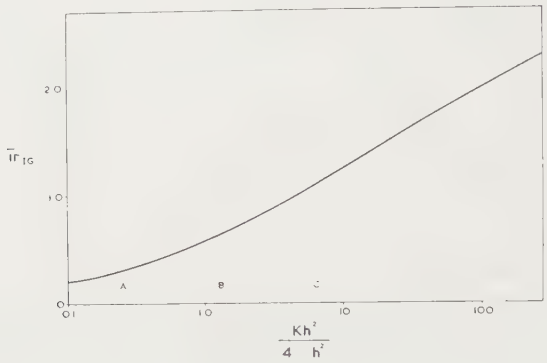


FIGURE 2

The Expected Genetic Superiority of Selected Sires (in Terms of σ_g in the Population) as a Function of $Kh^2/(4 - h^2)$

Tolerable Limits of Group Size

We have so far determined the group size which will give maximum superiority of chosen sires. It is useful to inquire further into the limits of variation in group size which can be tolerated before the genetic superiority falls off markedly. This can only be done empirically by drawing out curves of ΔG against n for different values of K and of h^2. Figures 3 shows the curves for $K = 100$ for different values of

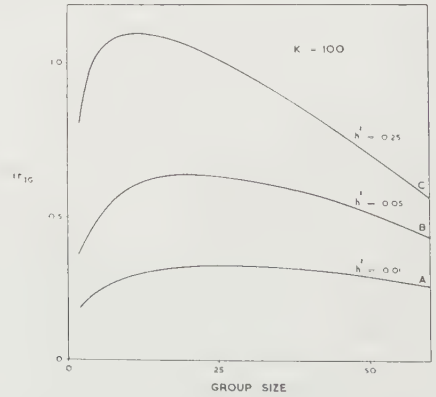


FIGURE 3

The Variation of Genetic Superiority of Selected Sires with Group Size ($K = 100$, h^2 Variable)

h^2. For simplicity, they have been drawn as continuous functions. They all must intersect the ΔG axis at $n = 0$ and $n = K$ (the latter implying $p = 1$, no selection). This imposes limitations upon the shape of the curves and on the optimum value of n. Thus in this case over the range of heritabilities from 0.01 to 0.25, ΔG will be within 10 percent of its maximum over the range of n from 10 to 26. For the comparison of the algebraic results with the actual curves, the three curves correspond to the points A, B, and C in Figures 1 and 2.

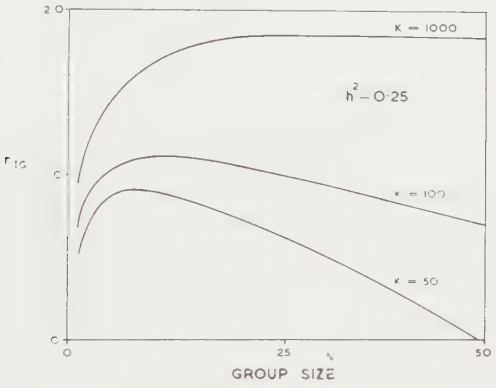


FIGURE 4

THE VARIATION OF GENETIC SUPERIORITY OF SELECTED SIRES WITH GROUP SIZE ($h^2 = 0.25$, K VARIABLE)

Figures 4 shows the set of curves for $h^2 = 0.25$ and different values of K. ΔG will be zero for all when $n = 0$, but the separate curves will again be zero at different points (when $n = K$, in fact).

There is thus not the same restraint in the position of the maximum. For $h^2 = 0.25$, there is a short range, in the neighborhood of $n = 12$, for which ΔG is within 10 percent of its maximum value over these ranges of K. But for lower values of h^2, these tolerable ranges of n no longer exist.

TABLE 2

THE RANGE OF GROUP SIZE GIVING GENETIC IMPROVEMENT WITHIN 10 PERCENT OF THE MAXIMUM FOR A HERITABILITY RANGE 0.01–0.25.

Testing ratio K	Group size
50	6–11
100	10–26
200	20–40
1000	60–90

Rather surprisingly, it seems that tolerable values of n are depending much more on the testing ratio K than on h^2. In fact, useful ranges of n can be given in terms of K which will give genetic superiority of chosen sires within 10 percent of the maximum value over the h^2 range 0.01 – 0.25. Table 2 gives these ranges for certain values of K. The preferred values within the range will increase as h^2 decreases.

The Effect of Non-Genetic Differences between Groups

The case which we have so far discussed assumed that the variance component between progeny groups was entirely genetic in origin. In many practical instances, however, this condition is not satisfied and there are environmental effects common to all members of a group. This reduces the effectiveness of selection as it introduces a non-genetic element into the between-group variance which cannot be reduced by an increase in group size. Formally we may describe this as a proportion (c^2) of the total variance, which is common to all members of a group. We have then the expression for r_{IG} .

$$r_{IG} = \sqrt{\frac{\frac{1}{4}nh^2}{1 + \frac{1}{4}(n-1)(h^2 + 4c^2)}}$$

$$= \sqrt{\frac{h^2}{h^2 + 4c^2}} \sqrt{\frac{\frac{1}{4}n(h^2 + 4c^2)}{1 + \frac{1}{4}(n-1)(h^2 + 4c^2)}}.$$

We may then generalise the presentation given above to include this case by writing

$$a' = \frac{4 - h^2 - 4c^2}{h^2 + 4c^2}$$

in the determination of optimum structure and by multiplying the expected improvement by the factor $\sqrt{h^2/h^2 + 4c^2}$. As a' will be less than a, it follows that for a given K the value of p at the optimum will be less than in the simpler case so that the optimum group size will be decreased and the number of sires tested increased. The generalizations we derived earlier about the minimum intensity of selection of 1 in 4 for such a programme will therefore still hold. The tolerable limits of group size will still hold, except that the critical range they refer to will be of $h^2 + 4c^2$ instead of h^2.

Examination of the expression for r_{IG} shows that it falls conceptually into two parts, if we introduce the idea of the breeding value of a sire under a repeatable but not measurable environmental bias. The first term, $\sqrt{h^2/(h^2 + 4c^2)}$, is then the correlation between the true breeding value and the actual value under that bias, and the second,

$$\sqrt{\frac{\frac{1}{4}n(h^2 + 4c^2)}{1 + \frac{1}{4}(n-1)(h^2 + 4c^2)}},$$

is the correlation between the biased value and its estimate, which has been subject to further error by sampling within the group.

Discussion

The treatment so far presented has been in terms of a progeny-testing programme. Exactly the same formulae will of course hold in half-sib family selection. In full-sib selection, the formula will have to be slightly modified. We should then write $a = (2 - h^2)/h^2$, rather than $(4 - h^2)/h^2$ in the half-sib case. We have presumed that the family size is variable at will. This of course is not so and the logical limitations in this are in some cases critical. Full-sib selection in cattle is obviously absurd and in poultry, full-sib family size is limited by the need to have only a limited hatching period. These factors, rather than the considerations presented here, will then be of overriding importance.

It may be of interest to illustrate this approach with two practical examples. The writer is associated with a progeny-testing scheme in dairy cattle at an artificial insemination centre. The essential fact of the scheme is the breeding of new bulls for testing from the best available proven sires. In each generation cycle, two sires ($S = 2$) will be selected on the basis of a total tested sample of 300 heifers ($N = 300$) so that $K = 150$. If we assume $h^2 = 0.25$ ($a = 15$) we find a K/a value of 10. Referring to Figures 1 and 2, we find that at the optimum, $p = 0.10$ and $G = 1.2\sigma_g$. Thus we should test 20 bulls on 15 daughters each—the actual intention is to test 12 on 25 daughters, because of the cost of buying and keeping too many sires unemployed. As σ_g is in the neighborhood of 80 gallons, the selected sires should show a genetic superiority of some 100 gallons, which in their daughters would be seen as a production superiority over their herd mates averaging 50 gallons.

Secondly, let us consider a flock of 500 hens to be improved by half-sib family selection, in which we take 5 selected sires as the lower limit possible. This gives $K = 100$ and, if we assume a heritability of egg production of 0.05, we have a K/a value of 1.25. At the optimum, this gives $p = 0.19$. We should therefore test 26 sires a year on 19 daughters each.

Summary

A general procedure is presented for the problem of optimum group size in progeny-testing and family selection programmes. This involves the testing ratio, K, the ratio of the number of individuals which can

be measured each generation to the number of sires (or family groups) that are to be selected. The optimum structure of the population and the probably genetic superiority of chosen groups are then dependent only on K, the heritability, h^2, and the genetic relationship within groups. The treatment has also been extended to cover problems arising from non-genetic differences between groups.

PARTITION OF EXPERIMENTAL VECTORS CONNECTED WITH MULTINOMIAL DISTRIBUTIONS

L. C. A. Corsten
Institute for Research on Varieties of Field Crops (I.V.R.O.)
Wageningen, The Netherlands

1. *Introduction*

The investigation of variates in contingency tables (which need not be orthogonal) often gives rise to a partition of the experimental result in components each illuminating one aspect of the problem in question. This is here expressed in terms of vectors. Compare Kuiper [9].

Our presentation is perhaps more transparent than previous papers dealing with similar subjects: Fisher [3], Fog [4], Irwin [6], Lancaster [10], [12], [13]. Moreover our method is general; hence it can be applied also to more intricate cases. Our results can be applied in genetics, which in fact motivated our research. This is why we give an introductory section (Sec. 2) on genetics.

2. *Genetics*

We consider externally perceptible properties of individuals (plants, animals), like the colour (green or yellow) of seeds, or of eyes (brown or blue) of men. Each property to be considered is determined by genes of only one locus, dominant A and recessive a, such that the only possible distinction with respect to that property is dominant (AA or Aa) or recessive (aa).

We assume that the choice of a paternal gamete by a zygote is stochastically independent of the choice of a maternal gamete. Moreover we assume equal viability for every combination of gametes.

Then it follows that crossing individuals Aa with each other will produce zygotes having the dominant and the recessive form of the property with probabilities $\frac{3}{4}$ and $\frac{1}{4}$ respectively.

We consider a second property and locus with genes B and b. If this locus is on a chromosome different from that carrying A and a, the dominance or recessivity will be independent for the two considered properties.

Crossing such individuals Aa/Bb with each other yields zygotes with probabilities as mentioned in the following table:

	A	a
B	$\frac{9}{16}$	$\frac{3}{16}$
b	$\frac{3}{16}$	$\frac{1}{16}$

(1)

where the capitals and lower case denote the dominant and recessive phenotypes respectively. The column and row totals represent the chances of the dominant and the recessive forms of the first and the second property respectively.

If both loci occur on the same chromosome and crossing-over does not take place, the gametes of an individual Aa/Bb are only AB and ab with equal probabilities. Crossing such individuals with each other produces zygotes for which the following table of probabilities appears:

	A	a
B	$\frac{3}{4}$	0
b	0	$\frac{1}{4}$

(2)

In a similar way crossing individuals Aa/bB with gametes Ab and aB yields:

	A	a
B	$\frac{1}{2}$	$\frac{1}{4}$
b	$\frac{1}{4}$	0

(2′)

In both (2) and (2′) the column and row totals are the same as in (1) and they have the same meaning.

If both loci occur on the same chromosome and crossing-over takes place, the gametes of an individual Aa/Bb (in the coupling phase) are AB, ab, aB and Ab with probabilities, say, $\frac{1}{2} - \frac{1}{2}W$, $\frac{1}{2} - \frac{1}{2}W$, $\frac{1}{2}W$, $\frac{1}{2}W$ respectively (with $0 \leq W \leq \frac{1}{2}$). Crossing such individuals with each other yields zygotes to which the following table of probabilities corresponds:

	A	a
B	$\frac{1}{4}(W^2 - 2W + 3)$	$\frac{1}{4}(-W^2 + 2W)$
b	$\frac{1}{4}(-W^2 + 2W)$	$\frac{1}{4}(W^2 - 2W + 1)$

(3)

In a similar way individuals Aa/bB (in the repulsion phase) where crossing-over takes place will have gametes AB, ab, aB, and Ab with

probabilities $\frac{1}{2}W$, $\frac{1}{2}W$, $\frac{1}{2} - \frac{1}{2}W$, and $\frac{1}{2} - \frac{1}{2}W$ respectively. The following table of probabilities for zygotes yielded by such gametes corresponds to this situation:

	A	a
B	$\frac{1}{4}(W^2 + 2)$	$\frac{1}{4}(1 - W^2)$
b	$\frac{1}{4}(1 - W^2)$	$\frac{1}{4}W^2$

(3′)

In both cases (3) and (3′), the column and row totals are the same as in (1), (2), and (2′). If $W = 0$ (crossing-over does not occur), (3) reduces to (2), and (3′) reduces to (2′): complete linkage. If $W = \frac{1}{2}$ both (3) and (3′) reduce to (1): the two properties are completely independent.

Increasing linkage (i.e. decreasing W) appears with respect to tables like (1) as an increase of the numbers in one diagonal and a decrease of the numbers in the other diagonal, row and column totals remaining constant; the contribution from linkage to each cell of the table has the same absolute value.

3. *Definition of components of a two by two table*

We consider a scheme of probabilities like (1), (a scheme with probabilities p_1 and q_1 for rows and p_2 and q_2 for columns and independence of these probabilities). This scheme (vector) is represented by

$$\begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix}, \quad \text{where} \quad p_i + q_i = 1 \qquad (i = 1, 2)$$

and it is called *the basis of the one-dimensional space of levels.*

It is our purpose to compare a scheme of counts arranged in a two by two table:

$$x = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \quad \text{where} \quad \sum_{i=1}^{4} x_i = n,$$

or rather of the relative frequencies x_i/n, with such a *level.* The experimental result may suggest that the independence is satisfied, that also the proportion $p_2 : q_2$ is not a bad description, but that the proportion $p_1 : q_1$ is wrong.

Then the following scheme may be more likely:

$$\begin{bmatrix} (p_1 + \alpha)p_2 & (p_1 + \alpha)q_2 \\ (q_1 - \alpha)p_2 & (q_1 - \alpha)q_2 \end{bmatrix}$$

wherein α is a suitably chosen real number. Also in this scheme the probabilities for the columns are p_2 and q_2 respectively and there is independence between row and column probabilities. The scheme is equal to

$$\begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix}.$$

We call

$$\alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix}$$

a *row effect* and

$$\begin{bmatrix} p_2 & q_2 \\ -p_2 & -q \end{bmatrix}$$

the *basis of the one-dimensional space of row effects.*
Analogously,

$$\begin{bmatrix} p_1 & -p_1 \\ q_1 & -q_1 \end{bmatrix}$$

is called the *basis of the one-dimensional space of column effects.*

It may also happen that the independence is satisfied, but that both proportions $p_1 : q_1$ and $p_2 : q_2$ are wrong, and that at first instance the following scheme seems to be more plausible:

$$\begin{bmatrix} (p_1 + \alpha)(p_2 + \beta) & (p_1 + \alpha)(q_2 - \beta) \\ (q_1 - \alpha)(p_2 + \beta) & (q_1 - \alpha)(q_2 - \beta) \end{bmatrix}. \tag{4}$$

The chances for the rows are $(p_1 + \alpha)$ and $(q_1 - \alpha)$ respectively, those for columns $(p_2 + \beta)$ and $(q_2 - \beta)$ respectively, the independence being maintained. But if, in accordance with the customary practice in the analysis of variance, we postulate additivity of row and column effects, the following vector appears:

$$\begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix} + \beta \begin{bmatrix} p_1 & -p_1 \\ q_1 & -q_1 \end{bmatrix} = \begin{bmatrix} p_1p_2 + \alpha p_2 + \beta p_1 & p_1q_2 + \alpha q_2 - \beta p_1 \\ q_1p_2 - \alpha p_2 + \beta q_1 & q_1q_2 - \alpha q_2 - \beta q_1 \end{bmatrix}. \tag{5}$$

In this scheme row and column totals are as desired, but the vector (4) is equal to (5) plus

$$\begin{bmatrix} \alpha\beta & -\alpha\beta \\ -\alpha\beta & \alpha\beta \end{bmatrix} = \alpha\beta \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The remark in Sec. 2 –last paragraph—may be expressed as follows: Table (3) can be written as the sum of (1) and

$$\frac{3 - 8W - 4W^2}{16} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

and Table (3′) as the sum of (1) and

$$\frac{4W^2 - 1}{16} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

So it is plausible to consider a vector

$$\gamma \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

as representative of linkage, or of disturbance of independence, or, in terms of the analysis of variance, as interaction. Therefore we call the vector

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

the *basis of the one-dimensional space of interactions.* We remark that in contrast with row and column effects, this basis does not depend on the form of the *level.*

In (4) and (5) we saw that, if both row and column effects are present, and moreover additive, a disturbance of independence with respect to the scheme (4) considered as *level* occurs. The disturbance is

$$\alpha\beta \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

If α and β are small, this *interaction* is very small and even absent if α and/or β are zero. Then this *interaction* is negligible in comparison with possible interaction from linkage.

With these remarks in mind it will be our problem to determine the coefficients in the following partition:

$$\begin{bmatrix} x_1/n & x_2/n \\ x_3/n & x_4/n \end{bmatrix} = \mu \begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix} + \beta \begin{bmatrix} p_1 & -p_1 \\ q_1 & -q_1 \end{bmatrix} + \gamma \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{6}$$

where μ, α, β, and γ are unknown.

4. *The Partition*

The conventional solution of μ, α, β, and γ is not simple, because the basis vectors on the righthand side of (6) are in general non-orthogonal. Orthogonality, i.e. ordinary inner products being zero, implies that the components can be obtained from orthogonal projections. Orthogonality can be obtained with a suitable choice of the following equivalent partition:

$$\begin{bmatrix} A\frac{x_1}{n} & B\frac{x_2}{n} \\ C\frac{x_3}{n} & D\frac{x_4}{n} \end{bmatrix} = \mu \begin{bmatrix} Ap_1p_2 & Bp_1q_2 \\ Cq_1p_2 & Dq_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} Ap_2 & Bq_2 \\ -Cp_2 & -Dq_2 \end{bmatrix} + \beta \begin{bmatrix} Ap_1 & -Bp_1 \\ Cq_1 & -Dq_1 \end{bmatrix} + \gamma \begin{bmatrix} A & -B \\ -C & D \end{bmatrix}. \tag{7}$$

The four components are orthogonal, if the 6 inner products, of which two turn out to be identical, vanish:

$$\begin{aligned} A^2p_1p_2^2 + B^2p_1q_2^2 - C^2q_1p_2^2 - D^2q_1q_2^2 &= 0 \\ A^2p_1^2p_2 - B^2p_1^2q_2 + C^2q_1^2p_2 - D^2q_1^2q_2 &= 0 \\ A^2p_1p_2 - B^2p_1q_2 - C^2q_1p_2 + D^2q_1q_2 &= 0 \\ A^2p_2 - B^2q_2 + C^2p_2 - D^2q_2 &= 0 \\ A^2p_1 + B^2p_1 - C^2q_1 - D^2q_1 &= 0. \end{aligned}$$

The third equation minus p_1 times the fourth equation, and the third equation minus p_2 times the fifth equation, yield $C^2 = (q_2/p_2)\,D^2$ and $B^2 = (q_1/p_1)D^2$ respectively. Substitution in the third equation yields: $A^2 = (q_1q_2/p_1p_2)D^2$. If we choose $D^2 = 1/q_1q_2$, we obtain :

$$A^2 = \frac{1}{p_1p_2}, \qquad B^2 = \frac{1}{p_1q_2}, \quad \text{and} \quad C^2 = \frac{1}{q_1p_2},$$

and these values happen to satisfy *all* equations.

Substituting the values for A, B, C, D in (7) we get:

$$\begin{bmatrix} \frac{x_1}{n\sqrt{p_1p_2}} & \frac{x_2}{n\sqrt{p_1q_2}} \\ \frac{x_3}{n\sqrt{q_1p_2}} & \frac{x_4}{n\sqrt{q_1q_2}} \end{bmatrix} = \mu \begin{bmatrix} \sqrt{p_1p_2} & \sqrt{p_1q_2} \\ \sqrt{q_1p_2} & \sqrt{q_1q_2} \end{bmatrix}$$

$$+\alpha \begin{bmatrix} \sqrt{\frac{p_2}{p_1}} & \sqrt{\frac{q_2}{p_1}} \\ -\sqrt{\frac{p_2}{q_1}} & -\sqrt{\frac{q_2}{q_1}} \end{bmatrix} + \beta \begin{bmatrix} \sqrt{\frac{p_1}{p_2}} & -\sqrt{\frac{p_1}{q_2}} \\ \sqrt{\frac{q_1}{p_2}} & -\sqrt{\frac{q_1}{q_2}} \end{bmatrix} + \gamma \begin{bmatrix} \frac{1}{\sqrt{p_1p_2}} & -\frac{1}{\sqrt{p_1q_2}} \\ -\frac{1}{\sqrt{q_1p_2}} & \frac{1}{\sqrt{q_1q_2}} \end{bmatrix}$$

or putting: $p_1p_2 = \pi_1$, $p_1q_2 = \pi_2$, $q_1p_2 = \pi_3$, and $q_1q_2 = \pi_4$, we obtain:

$$\begin{bmatrix} \frac{x_1}{n\sqrt{\pi_1}} & \frac{x_2}{n\sqrt{\pi_2}} \\ \frac{x_3}{n\sqrt{\pi_3}} & \frac{x_4}{n\sqrt{\pi_4}} \end{bmatrix} = \mu \begin{bmatrix} \sqrt{\pi_1} & \sqrt{\pi_2} \\ \sqrt{\pi_3} & \sqrt{\pi_4} \end{bmatrix} + \frac{\alpha}{\sqrt{p_1q_1}} \begin{bmatrix} \sqrt{\pi_3} & \sqrt{\pi_4} \\ -\sqrt{\pi_1} & -\sqrt{\pi_2} \end{bmatrix}$$

$$+ \frac{\beta}{\sqrt{p_2q_2}} \begin{bmatrix} \sqrt{\pi_2} & -\sqrt{\pi_1} \\ \sqrt{\pi_4} & -\sqrt{\pi_3} \end{bmatrix} + \frac{\gamma}{\sqrt{p_1q_1p_2q_2}} \begin{bmatrix} \sqrt{\pi_4} & -\sqrt{\pi_3} \\ -\sqrt{\pi_2} & \sqrt{\pi_1} \end{bmatrix}. \tag{8}$$

The four new vectors (in brackets) are perpendicular to each other, indeed ($\pi_1\pi_4 = \pi_2\pi_3$). They are moreover unit vectors. The required coefficients are thus simply the inner products of the vector [on the left hand side of (8)] to be projected, and the vector on which it is projected. For example, μ is obtained by taking the inner products of both sides of (8) with the unit vector of which μ is the coefficient. In this way we obtain:

$$\mu = \frac{1}{n}(x_1 + x_2 + x_3 + x_4) = 1,$$

$$\frac{\alpha}{\sqrt{p_1q_1}} = \frac{1}{n}\left[x_1\sqrt{\frac{\pi_3}{\pi_1}} + x_2\sqrt{\frac{\pi_4}{\pi_2}} - x_3\sqrt{\frac{\pi_1}{\pi_3}} - x_4\sqrt{\frac{\pi_2}{\pi_4}}\right]$$

$$= \frac{1}{n}\left[x_1\sqrt{\frac{q_1}{p_1}} + x_2\sqrt{\frac{q_1}{p_1}} - x_3\sqrt{\frac{p_1}{q_1}} - x_4\sqrt{\frac{p_1}{q_1}}\right]$$

$$= \frac{1}{n\sqrt{p_1q_1}}[(x_1 + x_2)q_1 - (x_3 + x_4)p_1],$$

or

$$\alpha = \frac{(x_1 + x_2)q_1 - (x_3 + x_4)p_1}{n}.$$

In a similar way:

$$\beta = \frac{(x_1 + x_3)q_2 - (x_2 + x_4)p_2}{n}$$

and

$$\gamma = \frac{x_1q_1q_2 - x_2q_1p_2 - x_3p_1q_2 + x_4p_1p_2}{n}.$$

Substituting these values in (8), subtracting the first term on the right from both sides and multiplying both sides by $\sqrt{n}$, we obtain:

$$\begin{aligned} X &= \begin{bmatrix} \frac{x_1 - n\pi_1}{\sqrt{n\pi_1}} & \frac{x_2 - n\pi_2}{\sqrt{n\pi_2}} \\ \frac{x_3 - n\pi_3}{\sqrt{n\pi_3}} & \frac{x_4 - n\pi_4}{\sqrt{n\pi_4}} \end{bmatrix} \\ &= \frac{[(x_1 + x_2)q_1 - (x_3 + x_4)p_1]}{\sqrt{np_1q_1}} \begin{bmatrix} \sqrt{\pi_3} & \sqrt{\pi_4} \\ -\sqrt{\pi_1} & -\sqrt{\pi_2} \end{bmatrix} \\ &+ \frac{[(x_1 + x_3)q_2 - (x_2 + x_4)p_2]}{\sqrt{np_2q_2}} \begin{bmatrix} \sqrt{\pi_2} & -\sqrt{\pi_1} \\ \sqrt{\pi_4} & -\sqrt{\pi_3} \end{bmatrix} \\ &+ \frac{(x_1q_1q_2 - x_2q_1p_2 - x_3p_1q_2 + x_4p_1p_2)}{\sqrt{np_1q_1p_2q_2}} \begin{bmatrix} \sqrt{\pi_4} & -\sqrt{\pi_3} \\ -\sqrt{\pi_2} & \sqrt{\pi_1} \end{bmatrix}. \end{aligned} \tag{9}$$

5. *Statistical considerations*

x has a multinomial probability distribution:

$$P(x) = \frac{n!}{x_1!x_2! \cdots x_k!} \pi_1^{x_1}\pi_2^{x_2} \cdots \pi_k^{x_k},$$

where

$$\sum_{i=1}^{k} \pi_i = 1 \quad \text{and} \quad \sum_{i=1}^{k} x_i = n,$$

or

$$P(x) = \frac{\frac{(n\pi_1)^{x_1}e^{-n\pi_1}}{x_1!} \times \frac{(n\pi_2)^{x_2}e^{-n\pi_2}}{x_2!} \times \cdots \times \frac{(n\pi_k)^{x_k}e^{-n\pi_k}}{x_k!}}{\frac{n^{\sum x_i}e^{-n}}{(\sum x_i)!}}. \tag{10}$$

In Sec. 4 we had the case $k = 4$ and $E(x_i) = n\pi_i$.

The same probability (10) will be obtained if we formally assert that the x_i have independent Poisson distributions with $\lambda_i = n\pi_i$ and probabilities $\lambda_i^{x_i} e^{-\lambda_i} / x_i!$ under the condition that their sum is equal to $\sum x_i$; for the sum of Poisson distributed variables has also a Poisson distribution with parameter $\sum \lambda_i$ (in this case equal to $\sum n\pi_i = n$), and its probability occurs in the denominator of $P(x)$, as is proper in the case of a conditional probability.

The same formal assertion implies that $X_i = (x_i - n\pi_i)/\sqrt{n\pi_i}$ has mean zero and unit variance. If further $n\pi_i$ is sufficiently large (e.g. > 9), the distribution of X_i will be approximated by the normal (0, 1) distribution.

The vector X that can be represented as a point in a k-dimensional space K, has thus approximately a probability density

$$Ce^{-X_1^2/2} \cdot e^{-X_2^2/2} \cdots = Ce^{-X^2/2}$$

however, with the restriction

$$\sum_{i=1}^{k} x_i = n, \quad \text{or} \quad \sum_{i=1}^{k} X_i \sqrt{n\pi_i} + \sum_{i=1}^{k} n\pi_i = n, \quad \text{or} \quad \sum_{i=1}^{k} X_i \sqrt{\pi_i} = 0.$$

In other words X is situated in the $(k-1)$ dimensional subspace of K, perpendicular to the vector $(\sqrt{\pi_1}, \sqrt{\pi_2}, \cdots, \sqrt{\pi_k})$.
Thus we find the (known) result that

$$X^2 = \sum_{i=1}^{k} \frac{(x_i - n\pi_i)^2}{n\pi_i}$$

has approximately a χ^2-distribution with $(k-1)$ dimensions (or degrees of freedom).

In Sec. 4 we succeeded in splitting up X (situated in a 3-dimensional space) in three perpendicular components each of which has a special meaning. From the foregoing it follows that, if $E(x_i) = n\pi_i$, the square of every component vector—which is equal to the square of the coefficient belonging to it in (9)—has approximately a χ^2-distribution with one dimension.

In the light of our definitions in Sec. 3, we can even say that the square of the length of the first component has a one-dimensional χ^2-distribution, if only row effect is absent irrespective of whether the other effects are present or not. The same is true for the second component if only column effect is absent, and the same for the third if there are no interaction and not too large simultaneous main effects. In other words, each of these components affords us a specific test criterion for the null hypothesis that there is no row effect, or no column

effect, or no interaction which can be tested independent of the validity of the other two hypotheses to a certain extent. The last restriction concerns the simultaneous occurrence of considerable row and column effects which influences the component for interaction. The numerical value of each of these three statistics is thus obtained by splitting up the known test criterion X^2 for goodness of fit in three terms, the first of which is equal to

$$\frac{[(x_1 + x_2)q_1 - (x_3 + x_4)p_1]^2}{np_1q_1},$$

the second equal to

$$\frac{[(x_1 + x_3)q_2 - (x_2 + x_4)p_2]^2}{np_2q_2},$$

and the third is the rest:

$$\frac{[x_1q_1q_2 - x_2q_1q_2 - x_3p_1q_2 + x_4p_1p_2]^2}{np_1q_1p_2q_2}.$$

Just as is the case in tests with X^2, the critical region in the one-dimensional χ^2-distributions will be a one-sided upper critical region.

Example (Fisher [3]). Counting of seedlings of self-fertilized maize, which was heterozygous Aa/bB (i.e. in repulsion phase) with respect to two properties, viz. starchy versus sugary and green versus white, gave the following results:

	starchy	sugary
green	1997	904
white	906	32

If the properties are independent, then the vector *level* will be equal to

$$\begin{bmatrix} \frac{9}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{1}{16} \end{bmatrix} \quad \text{with} \quad p_1 = p_2 = \tfrac{3}{4}.$$

The ordinary χ^2-test criterion for these probabilities will be found to be equal to $12.21 + 47.13 + 180.19 = 287.69$ which is significant, e.g. at the 1% level.

In order to investigate the origin of this discrepancy from our

expectation (compare case (3') in Sec. 2), we calculate the squares of the coefficients in (9): corresponding with row effect

$$\frac{[\frac{1}{4}(x_1 + x_2) - \frac{3}{4}(x_3 + x_4)]^2}{n\frac{1}{4}\cdot\frac{3}{4}} = \frac{[x_1 + x_2 - 3(x_3 + x_4)]^2}{3n} = 0.65,$$

with column effect

$$\frac{[x_1 + x_3 - 3(x_2 + x_4)]^2}{3n} = 0.78,$$

with interaction

$$\frac{(x_1 - 3x_2 - 3x_3 + 9x_4)^2}{9n} = 286.27,$$

which are the same as those obtained by Fisher. The sum of these squares is indeed equal to the value of X^2. Further, we see that the deviation from our expectation is practically exclusively due to interaction or linkage.

In many cases we have no theoretical indications about the chances for rows and columns. Then (and in the case that row and column effects appear to be present and a further investigation of dependence is wished), the unknown chances are estimated according to the maximum likelihood method assuming independence between row and column chances. These estimates are proportional to the marginal totals and imply that X^2 has a one-dimensional χ^2-distribution.

A more elementary way of approaching this problem and its consequence for X^2 and for the partition of X^2 runs as follows. Consider the set of all possible values of x/n in a 2×2 table with the independent chances p_1 and p_2 unknown and with fixed n. In the subset where the marginal totals are fixed, the conditional probability distribution of x/n will be:

$$\frac{\dfrac{n!}{x_1!x_2!x_3!x_4!}(p_1p_2)^{x_1}(p_1q_2)^{x_2}(q_1p_2)^{x_3}(q_1q_2)^{x_4}}{\dfrac{n!}{(x_1 + x_2)!(x_3 + x_4)!}p_1^{x_1+x_2}q_1^{x_3+x_4}\dfrac{n!}{(x_1 + x_3)!(x_2 + x_4)!}p_2^{x_1+x_3}q_2^{x_2+x_4}}$$

which appears to be independent of the unknown p_i. We may therefore assume any p_i to be true in the multinomial distribution from which the conditional distribution in the chosen subset can be obtained. It is thus permissible to assume that in the original multinomial distribution the p_i are equal to the chosen marginal totals. We saw that the unconditional distribution of X *with the assumed* p_i *as parameters*

could be approximated by the three-dimensional normal distribution. By imposing the conditions of the considered subset to this X, i.e. that the marginal totals of x/n should be equal to the assumed row and column probabilities, X will be limited to a one-dimensional space. In the subset the conditional distribution of such a X^2 will be that of a one-dimensional χ^2. Because this conditional distribution does not depend on the marginal totals defining the subset, it is valid in general. Further, it follows that by assuming the p_i equal to the corresponding observed frequencies, the two components for row and column effects in the partition (9) of such a X^2 will vanish, and that by this very choice the square of the component for interaction has a one-dimensional χ^2-distribution.

In this particular case the component for interaction can be reduced by substituting the marginal totals for p_i and q_i to

$$\frac{n(x_1x_4 - x_2x_3)^2}{(x_1 + x_2)(x_3 + x_4)(x_1 + x_3)(x_2 + x_4)}$$

or using: $x_1x_4 - x_2x_3 = x_1(x_1 + x_2 + x_3 + x_4) - (x_1 + x_2)(x_1 + x_3)$ to

$$\frac{\left[x_1 - \frac{(x_1 + x_2)(x_1 + x_3)}{n}\right]^2}{n\left(\frac{x_1 + x_2}{n}\right)\left(\frac{x_1 + x_3}{n}\right)\left(\frac{x_3 + x_4}{n}\right)\left(\frac{x_2 + x_4}{n}\right)}$$

which will be useful later on.

Example: If we take the same numbers as in the previous example and we suppose nothing to be known about probabilities, the estimates of expected numbers are

$$\begin{bmatrix} 2193.69 & 707.30 \\ 709.30 & 228.69 \end{bmatrix}$$

and $X^2 = 17.64 + 54.70 + 54.55 + 169.17 = 296.06$, which is much more than the 0.1% point of the one-dimensional χ^2-distribution (10.827). So there is interaction or lack of independence. The fact that the latter X^2 turns out to be larger than the former should not surprise us, as maximum likelihood estimators do not in general lead to a minimal value of χ^2. Therefore it is possible that in this example the values $p_1 = p_2 = \frac{3}{4}$ yield a lower value of the goodness of fit criterion X^2 than the values of p_1 and p_2, estimated by the maximum likelihood method, do.

6. 2^n *tables*

It is not difficult to generalize the treatment of 2×2 tables to that of 2^n tables. We indicate the generalization by presenting the case $n = 3$. In that case the basis vector for the space of *levels* is defined to be a three-way table with given probabilities for the completely independent rows, columns, and layers ($p_i + q_i = 1$). The probabilities for each of the eight cells are then given by Fig. 1. The probability for row 1 (back face) is p_1 , for row 2 (front face) q_1 , for column 1 (left side) p_2 , for column 2 (right side) q_2 , for layer 1 (bottom) p_3 , and for layer 2 (upper face) q_3 .

We call a vector *row effect* if the sum of this vector and the vector *level* looks like Fig. 2. The probability for back space is $p_1 + \alpha$, for front face $(q_1 - \alpha)$, for left side p_2 , for right side q_2 , for bottom p_3 , for upper face q_3 , all these being independent.

We choose as a basis vector for *row effects*, *column effects*, and *layer effects* the vectors represented by Figs. 3, 4, and 5 respectively.

We call a vector *row* $\times$ *column interaction*, if the sum of this vector and the vector *level* looks like Fig. 6. The probabilities in the six faces are equal to those in the vector *level* by itself, so that there is no *main effect* in this sum. The chances in the vertical edges are:

$$\begin{bmatrix} p_1p_2 + \delta & p_1q_2 - \delta \\ q_1p_2 - \delta & q_1q_2 + \delta \end{bmatrix}$$

so that independence between rows and columns is disturbed. Further, the probabilities for layers are independent of those for the other classifications.

We choose as a basis vector for *row* $\times$ *column interaction*, *row* $\times$ *layer interaction*, and *column* $\times$ *layer interaction* the vectors represented by Figs. 7, 8, and 9 respectively.

Finally, we call a vector *second-order interaction* if the sum of this vector and the vector *level* looks like Fig. 10.

The probabilities in the six faces are equal to those in the vector *level* by itself, so that *main effect* is absent in this sum. The probabilities in the twelve edges are also equal to those in the vector *level* by itself, so that there is no *first-order interaction* in the sum.

But the disturbance of independence between row and column probabilities in the bottom is different from that in the upper face. Similar remarks can be made about row and layer probabilities and about column and layer probabilities. In other words, the relation between any pair of classifications cannot be described without involving the third classification.

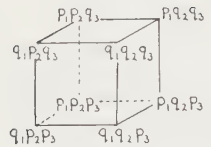


Fig. 1
Level

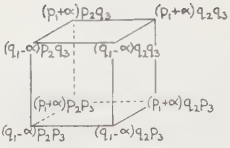


Fig. 2
Level + row effect

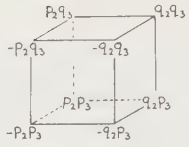


Fig. 3
Basis for row effects

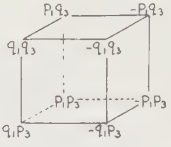


Fig. 4
Basis for column effects

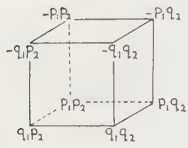


Fig. 5
Basis for layer effects

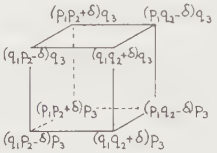


Fig. 6
Level + row × column interaction

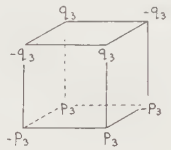


Fig. 7
Basis for row × column interaction

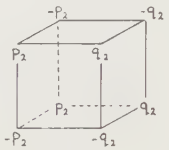


Fig. 8
Basis for row × layer interaction

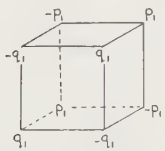


Fig. 9
Basis for column × layer interaction

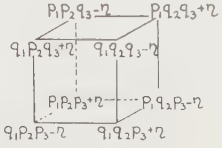


Fig. 10
Level + second order interaction

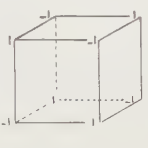


Fig. 11
Basis for second order interaction

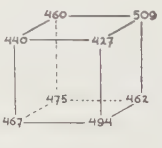


Fig. 12
Example of a 2^3 table (*Roberts* e.a.)

We choose as a basis vector for *second-order interaction* the vector represented by Fig. 11.

About the thus defined interactions, remarks similar to that in Sec. 3 may be made. If main effects and interactions are present, we have in the place of $p_1p_2p_3$ a chance equal to

$$(p_1 + \alpha)(p_2 + \beta)(p_3 + \gamma) + \delta(p_3 + \gamma) + \epsilon(p_2 + \beta) + \zeta(p_1 + \alpha) + \eta$$
$$= p_1p_2p_3 + \alpha p_2p_3 + \beta p_1p_3 + \gamma p_1p_2 + (\alpha\beta + \delta)p_3 + (\alpha\gamma + \epsilon)p_2$$
$$+ (\beta\gamma + \zeta)p_1 + (\alpha\zeta + \beta\epsilon + \gamma\delta + \alpha\beta\gamma + \eta).$$

From this it will be seen that three main effects together introduce contributions to all of the four interactions. In that case certain first-order interactions may nullify the additional contribution to the second-order interaction. If there are two main effects only, there will be a contribution to the first-order interaction between them. If in this case one or both of the other first-order interactions are non-

zero, a contribution to the second-order interaction may take place. If there is only one main effect and simultaneously an interaction between the two other classifications, there will be a contribution to the second-order interaction. If there are no main effects at all, the second-order interaction obtains no contributions from possible first-order interactions. We may conclude that disturbing contributions to interactions may be caused by main effects, but that they will be negligible if α, β, and γ are small. The drawback of such contributions is not great as it is not our intention to estimate the different effects, but to investigate the origin of discrepancies between expectation and experimental result.

It will be our purpose to split up a vector of experimental numbers x_i/n $(i = 1, 2, \cdots, 8;\ \sum_i x_i = n)$ into eight components in the directions of the defined basis vectors. To facilitate this partition, we again divide the numbers on similar places by the square root of the probability of that place in the vector *level*. Putting those probabilities equal to π_i, we obtain the following formulation of our problem:

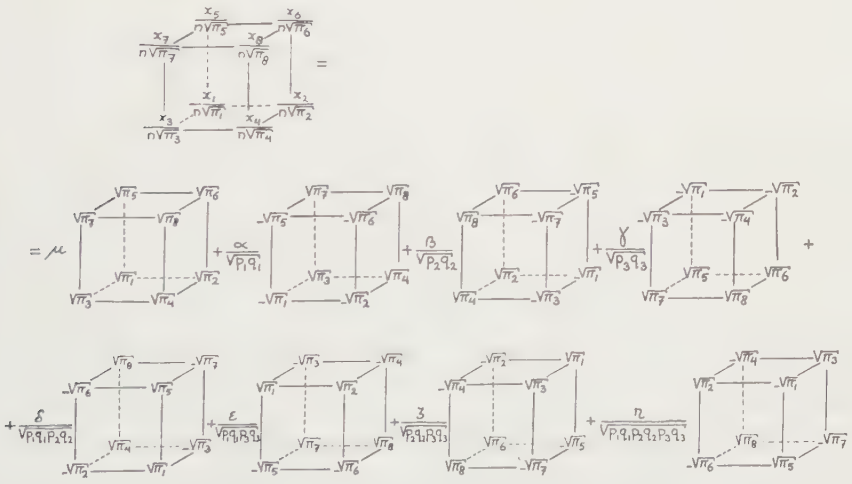


Fig. 13
Equation (11): Partition of 2^3 table

Calculate the coefficients μ, α, β, $\cdots$ η in the equation represented by Fig. 13 which will be called equation (11). All the "cubes" on the right-hand side of this equation are orthogonal unit vectors.

$$\left(\text{Observe that } \frac{\pi_1}{\pi_2} = \frac{\pi_3}{\pi_4} = \frac{\pi_5}{\pi_6} = \frac{\pi_7}{\pi_8} \quad \text{and} \quad \frac{\pi_1}{\pi_3} = \frac{\pi_2}{\pi_4} = \frac{\pi_5}{\pi_7} = \frac{\pi_6}{\pi_8}.\right)$$

The required coefficients are thus again the inner products of the vector on the left side to be projected and the vector on which it is projected:

$$\mu = 1,$$

$$\alpha = \frac{(x_1 + x_2 + x_5 + x_6)q_1 - (x_3 + x_4 + x_7 + x_8)p_1}{n},$$

$$\beta = \frac{(x_1 + x_3 + x_5 + x_7)q_2 - (x_2 + x_4 + x_6 + x_8)p_2}{n},$$

$$\gamma = \frac{(x_1 + x_2 + x_3 + x_4)q_3 - (x_5 + x_6 + x_7 + x_8)p_3}{n},$$

$$\delta = \frac{(x_1 + x_5)q_1q_2 - (x_2 + x_6)q_1p_2 - (x_3 + x_7)p_1p_2 + (x_4 + x_8)p_1p_2}{n},$$

$$\epsilon = \frac{(x_1 + x_2)q_1q_3 - (x_3 + x_4)p_1q_3 - (x_5 + x_6)q_1p_3 + (x_7 + x_8)p_1p_3}{n},$$

$$\zeta = \frac{(x_1 + x_3)q_2q_3 - (x_2 + x_4)p_2q_3 - (x_5 + x_7)q_2p_3 + (x_6 + x_8)p_2p_3}{n},$$

$$\eta = \frac{x_1q_1q_2q_3 - x_2q_1p_2q_3 - x_3p_1q_2q_3 + x_4p_1p_2q_3}{n}$$

$$- \frac{x_5q_1q_2p_3 - x_6q_1p_2p_3 - x_7p_1q_2p_3 + x_8p_1p_2p_3}{n}.$$

Substituting these values in equation (11), Fig. 13, subtracting the first term on the fight from both sides and multiplying on both sides by $\sqrt{n}$, we get on the left a vector X consisting of numbers $(x_i - n\pi_i)/\sqrt{n\pi_i}$ the square of which is again equal to the well-known goodness of fit criterion. X^2 has approximately a χ^2-distribution with 7 degrees of freedom, if the null hypothesis that the vector *level* contains the probabilities for each cell is true. Further, the seven components of X on the right are perpendicular, and X^2 is equal to the sum of the seven squares of these components. Each of these squares has under the null hypothesis a χ^2-distribution with one degree of freedom.

Now again every component can serve as a statistic for testing a specific hypothesis. With the first three components we test $\alpha = 0$, $\beta = 0$, and $\gamma = 0$ respectively, with the following three components, the hypotheses $\delta = 0$, $\epsilon = 0$, and $\zeta = 0$, under the condition that $\alpha\beta$, $\alpha\gamma$, and $\beta\gamma$ are negligible respectively, and with the last component, the hypothesis $\eta = 0$, under the condition that $\alpha\zeta + \beta\epsilon + \gamma\delta + \alpha\beta\gamma$ is

negligible. Each of these hypotheses is tested on the assumption of independence of the three classifications.

The mentioned restrictions do not trouble us if no theoretical chances are available. Generalizing the considerations at the end of Sec. 5, we take the p_i equal to marginal totals with the consequence that $\alpha = \beta = \gamma = 0$. If any product $\alpha\beta$, $\alpha\gamma$, or $\beta\gamma$ appears to be not negligible and a further investigation of interactions is required, it is recommended to take the relative p_i from marginal totals.

A more difficult situation will arise if a first-order interaction appears to be considerable or if it is expected in advance that any pair of classifications is not independent. This situation is obviously contrary to the assumption of independence of the three classifications and will be considered in Sec. 8.

Example: (Roberts, Dawson, and Madden [14].) Crossing mice $AaBbCc$ with $aabbcc$ gave numbers represented in Fig. 12, (A in back face, B in left side, and C in bottom). In the level p_i was $\frac{1}{2}$ ($i = 1, 2, 3$).

The value of χ^2 for row $\times$ layer interaction, e.g. is

$$\frac{\epsilon^2}{np_1q_1p_3q_3} = \frac{(x_1 + x_2 + x_7 + x_8 - x_3 - x_4 - x_5 - x_6)^2}{n},$$

and for second-order interaction is

$$\frac{\eta^2}{np_1q_1p_2q_2p_3q_3} = \frac{(x_1 + x_4 + x_6 + x_7 - x_2 - x_3 - x_5 - x_8)^2}{n}.$$

The seven values for χ^2 are:

for row effect	1.63
column effect	0.67
layer effect	1.03
row $\times$ column interaction	0.13
row $\times$ layer interaction	4.25
column $\times$ layer interaction	0.13
second-order interaction	2.79
total	10.63

The one-dimensional χ^2 has at the 5% level of significance the critical value: 3.84; the 7-dimensional χ^2: 14.07. The total χ^2 is not significant. Concluding that the row $\times$ layer interaction is significant, would be rash: the probability that at least one of seven one-dimensional χ^2 is larger than 4.25 is about 0.24. So there is only a slight indication that linkage between A and C may exist.

It is not always possible to attach a simple meaning to a second-order interaction, but in cases like this it could be caused by the fact

that one of the eight gene combinations, as a consequence of diminished viability, is much less frequent than is expected on account of main effects and first-order interactions only.

About the calculations we can remark the following: χ^2 for row $\times$ column interaction can be determined as the interaction χ^2 in the two by two row-column table which can be formed by adding along the four vertical edges; the formula for δ is then the same as that for γ in a two by two table. Thus one can calculate this χ^2 again from the test criterion X^2 for this two by two table and by subtraction of the main effects for rows and columns. The second-order interaction χ^2 is found as the difference of the test criterion X^2 for the 2^3 table and the sum of those for main effects and first-order interactions.

If the probabilities p_i are not known, we take them such that the main effects equal zero. In our example we then obtain the value χ^2 for

row $\times$ column interaction	0.12
row $\times$ layer interaction	4.34
column $\times$ layer interaction	0.14
second-order interaction	2.69
total	7.29

The 4-dimensional χ^2 has at the 5% level of signicance a critical value: 9.49. The total χ^2 is not significant. The probability that at least one of four one-dimensional χ^2 is larger than 4.34 is about 0.14; so there is a slight indication that linkage between A and C may exist.

An example of the partition of a 2^5 table in 31 components is given by Haldane [5].

7. $m \times n \times \cdots$ *tables*

In some particular cases a partition of $m \times n \times \cdots$ tables (and even of non-orthogonal tables) may have sense. We indicate the case of a 2×3 table which can be generalized easily.

An inquiry into the attitude with respect to a political proposal may be summarized in a table of experimental counts:

	for	against	no opinion
men	x_1	x_2	x_3
women	x_4	x_5	x_6

In this case an appropriate definition of a vector *level* will be:

$$\begin{bmatrix} p_1p_2p_3 & p_1p_2q_3 & p_1q_2 \\ q_1p_2p_3 & q_1p_2q_3 & q_1q_2 \end{bmatrix}, \quad \text{where} \quad p_i + q_i = 1, \qquad (i = 1, 2, 3).$$

The chances for rows (men and women) are p_1 and q_1 respectively, those for columns 1 and 2 together (politically interested), and for column 3 (politically not interested), p_2 and q_2 respectively, and those of columns 1 and 2, p_2p_3 and p_2q_3 (i.e. p_3 and q_3 under the condition of being in column 1 or 2). Moreover, all these probabilities are independent. Without further explanation we define as *basis vector* for *row effect*:

$$\begin{bmatrix} p_2p_3 & p_2q_3 & q_2 \\ -p_2p_3 & -p_2q_3 & -q_2 \end{bmatrix}.$$

Column effects will have two basis vectors, one corresponding with modifications of p_2 and q_2, independence being maintained:

$$\begin{bmatrix} p_1p_3 & p_1q_3 & -p_1 \\ q_1p_3 & q_1q_3 & -q_1 \end{bmatrix}, \tag{12}$$

and one corresponding with similar modifications of p_3 and q_3 :

$$\begin{bmatrix} p_1p_2 & -p_1p_2 & 0 \\ q_1p_2 & -q_1p_2 & 0 \end{bmatrix} \text{ or rather } \begin{bmatrix} p_1 & -p_1 & 0 \\ q_1 & -q_1 & 0 \end{bmatrix}. \tag{13}$$

If two such column effects occur together, viz. β times (12) and γ times (13) and, if they are additive, we obtain:

$$\begin{bmatrix} p_1(p_2+\beta)\left(p_3+\frac{\gamma}{p_2}\right)-\frac{\beta\gamma p_1}{p_2} & p_1(p_2+\beta)\left(q_3-\frac{\gamma}{p_2}\right)+\frac{\beta\gamma p_1}{p_2} & p_1(q_2-\beta) \\ q_1(p_2+\beta)\left(p_3+\frac{\gamma}{p_2}\right)+\frac{\beta\gamma q_1}{p_2} & q_1(p_2+\beta)\left(q_3-\frac{\gamma}{p_2}\right)-\frac{\beta\gamma q_1}{p_2} & q_1(q_2-\beta) \end{bmatrix}.$$

This is:

$$\frac{\beta\gamma}{p_2}\begin{bmatrix} p_1 & -p_1 & 0 \\ q_1 & -q_1 & 0 \end{bmatrix}$$

less than a level with probabilities p_1 , $p_2 + \beta$, $p_3 + \gamma/p_2$, etc. With respect to this level there is a deficit of $\beta\gamma/p_2$ times the column effect (13). As β will be small in comparison with p_2 , $\beta\gamma/p_2$, however, will be negligible in comparison with γ, and even will be zero, if β or γ is zero. Similar remarks can be made about simultaneous occurrence of

additive row and column effects; there will be disturbances of independence which will be negligible in comparison with proper interactions. We choose as *basis vectors for interactions*:

$$\begin{bmatrix} p_3 & q_3 & -1 \\ -p_3 & -q_3 & 1 \end{bmatrix}$$

which represents the disturbance of independence between row probabilities, p_1 and q_1 , and column probabilities, p_2 and q_2 , and

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

which represents a similar disturbance of independence between row probabilities and column probabilities, p_3 and q_3 . It will be our purpose again to split up a vector of experimental numbers x_i/n ($i = 1, 2, \cdots 6$; $\sum x_i = n$) in six components in the directions of the defined basis vectors. We divide again the numbers on similar places by the square root of the probability of that place in the vector *level*. Putting those probabilities equal to π_i , we obtain:

$$\begin{aligned}
\begin{bmatrix} \frac{x_1}{n\sqrt{\pi_1}} & \frac{x_2}{n\sqrt{\pi_2}} & \frac{x_3}{n\sqrt{\pi_3}} \\ \frac{x_4}{n\sqrt{\pi_4}} & \frac{x_5}{n\sqrt{\pi_5}} & \frac{x_6}{n\sqrt{\pi_6}} \end{bmatrix} &= \mu \begin{bmatrix} \sqrt{\pi_1} & \sqrt{\pi_2} & \sqrt{\pi_3} \\ \sqrt{\pi_4} & \sqrt{\pi_5} & \sqrt{\pi_6} \end{bmatrix} \\
&+ \frac{\alpha}{\sqrt{p_1 q_1}} \begin{bmatrix} \sqrt{\pi_4} & \sqrt{\pi_5} & \sqrt{\pi_6} \\ -\sqrt{\pi_1} & -\sqrt{\pi_2} & -\sqrt{\pi_3} \end{bmatrix} \\
&+ \frac{\beta}{\sqrt{p_2 q_2}} \begin{bmatrix} \sqrt{p_1 q_2 p_3} & \sqrt{p_1 q_2 q_3} & -\sqrt{p_1 p_2} \\ \sqrt{q_1 q_2 p_3} & \sqrt{q_1 q_2 q_3} & -\sqrt{q_1 p_2} \end{bmatrix} \\
&+ \frac{\gamma}{\sqrt{p_2 p_3 q_3}} \begin{bmatrix} \sqrt{p_1 q_3} & -\sqrt{p_1 p_3} & 0 \\ \sqrt{q_1 q_3} & -\sqrt{q_1 p_3} & 0 \end{bmatrix} \\
&+ \frac{\delta}{\sqrt{p_1 q_1 p_2 q_2}} \begin{bmatrix} \sqrt{q_1 q_2 p_3} & \sqrt{q_1 q_2 q_3} & -\sqrt{q_1 p_2} \\ -\sqrt{p_1 q_2 p_3} & -\sqrt{p_1 q_2 q_3} & \sqrt{p_1 p_2} \end{bmatrix} \\
&+ \frac{\epsilon}{\sqrt{p_1 q_1 p_2 p_3 q_3}} \begin{bmatrix} \sqrt{q_1 q_3} & -\sqrt{q_1 p_3} & 0 \\ -\sqrt{p_1 q_3} & \sqrt{p_1 p_3} & 0 \end{bmatrix}.
\end{aligned} \tag{14}$$

The six new vectors (in brackets) on the right are orthogonal unit vectors. The coefficients are obtained again by taking inner products:

$$\mu = 1,$$

$$\alpha = \frac{1}{n}[(x_1 + x_2 + x_3)q_1 - (x_4 + x_5 + x_6)p_1],$$

$$\beta = \frac{1}{n}[(x_1 + x_2 + x_4 + x_5)q_2 - (x_3 + x_6)p_2],$$

$$\gamma = \frac{1}{n}[(x_1 + x_4)q_3 - (x_2 + x_5)p_3 ,$$

$$\delta = \frac{1}{n}[(x_1 + x_2)q_1q_2 - x_3q_1p_2 - (x_4 + x_5)p_1q_2 + x_6p_1p_2],$$

$$\epsilon = \frac{1}{n}[x_1q_1q_3 - x_2q_1p_3 - x_4p_1q_3 + x_5p_1p_3].$$

Substituting these values in equation (14), subtracting the first term on the right from both sides and multiplying on both sides by $\sqrt{n}$, we get on the left a vector X consisting of numbers $(x_i - n\pi_i)/\sqrt{n\pi_i}$, the square of which is again equal to the goodness of fit criterion. Under the null hypothesis that the vector *level* is true, the square of each term on the right has a χ^2-distribution with one degree of freedom. It will be remarked that the χ^2 for row effect, the first column effect, and the first interaction can be obtained as the χ^2 for the similar components in a 2×2 table which is deduced from our table by amalgamating the columns 1 and 2. The χ^2 for the second column effect and the second interaction are equal to

$$\frac{[(x_1 + x_4)q_3 - (x_2 + x_5)p_3]^2}{p_3q_3 \cdot np_2} \quad \text{and} \quad \frac{(x_1q_1q_3 - x_2q_1p_3 - x_4p_1q_3 + x_5p_1p_3)^2}{p_1q_1p_3q_3 \cdot np_2}$$

respectively. They will be obtained as the χ^2 for column effect and interaction in the 2×2 table consisting of the first and second column [see equation (9)], with the restriction, however, that we do not use the experimental total, $x_1 + x_2 + x_4 + x_5$, in the denominator, but the expected total np_2.

If there are no theoretical values p_i, then we take them such that main effects are absent, i.e. from marginal totals (these are also maximum likelihood estimations). Then the χ^2 for the first interaction will be found as the ordinary χ^2 test criterion for the 2×2 table obtained by amalgamating the first two columns and the χ^2 for the second interaction as the χ^2 *for interaction* in a 2×2 table consisting of the first

two columns, but with row and column probabilities estimated from the marginal totals of the whole 2×3 table. In this case np_2 is of course taken equal to $x_1 + x_2 + x_4 + x_5$.

Example: The inquiry mentioned in the beginning of this section may have the following result:

	for	against	no opinion
men	1154	475	243
women	1083	442	362

(These numbers are taken from *Introduction to the Theory of Statistics* by A. M. Mood, page 273, where they occur as an example of a 2×3 contingency table.) Theoretical probabilities being absent, we calculate the two-dimensional χ^2 for the whole table with the aid of expected values obtained from marginal totals:

1114.04	456.67	301.29
1122.96	460.33	303.71

and we find 26.78. The critical value of a two-dimensional χ^2 is 9.21 at the 1% level. So we conclude that there is association in our numbers. In order to investigate the origin of this association, we calculate the χ^2 for the first interaction, which is found as the ordinary χ^2 from the amalgamated table:

	opinion	no opinion
men	1629	243
women	1525	362

and with expected values also obtained by amalgamation of the calculated expected values:

$$\begin{bmatrix} 1570.71 & 301.29 \\ 1583.29 & 303.71 \end{bmatrix}$$

and we find 26.77. By subtraction it follows that χ^2 for the second interaction is equal to 0.01. We conclude therefore that there was a difference in interest between the sexes, but that no difference in attitude against the proposal could be detected between men and women.

In general each one-dimensional χ^2 in a contingency table, where the probabilities for rows and columns are estimated from the data and where a similar partition takes place, is connected with one of the 2×2 tables which are obtained by successive amalgamations of the data. Such a χ^2 is calculated as the ordinary χ^2 for *interaction* (as derived in Sec. 5) in this 2×2 table of which the table of expected values is obtained by a similar amalgamation of the *complete* scheme of expected values (estimated by the maximum likelihood method). The last results where theoretical chances are unknown are the same as those of Lancaster [10] and Kimball [8]. In connection with the foregoing it may be remarked that the method, suggested by Lancaster [11] as an exact one, for calculating χ^2 in a contingency table where cells with small expectations are pooled, does not seem to be correct.

When pooling of cells takes place it is not correct to say only that one or more one-dimensional χ^2 reduced to zero. If, e.g. in a 3×3 table, two cells in the same row are pooled, in other words are conceived as one cell, these cells do not contribute information about the estimation of the probabilities for the columns to which these cells belonged before pooling. For the vector level can be described as:

$p_1p_2p_3$	$p_1p_2q_3p_4$	$p_1p_2q_3q_4$
$p_1q_2p_3$	$p_1q_2q_3p_4$	$p_1q_2q_3q_4$
q_1p_3	q_1q_3	

with $p_i + q_i = 1$ $(i = 1, 2, 3, 4)$.

The estimation of p_1 and p_2 takes place from row totals in the same way as in the 3×3 table. Similarly, the p_3 will be estimated from the totals of the first column and the total of the second and the third column together. But p_4 will be estimated from the proportion of the totals in what has been left of the second and the third column, namely in the first and the second row. In the expressions for several of the one-dimensional chi-squares that do not vanish by pooling, the (estimated) values of p_4 and q_4 will in general be different from the values obtained by estimation in the complete 3×3 table.

The same conclusion follows from the fact that as a consequence of such pooling in the complete partition of a 3×3 table according to main effects and interactions, the basis of column effect with respect to p_4 will coincide with that for interaction in the 2×2 table formed by the second and third column on the one hand, and by the sum of

the first and second row and the third row on the other hand. Thus not only this interaction but also the mentioned main effect should vanish in the partition after pooling, in order that X^2, which measures the discrepancy with respect to a level, as represented here, contain three interaction components only.

8. *Further remarks about* 2^3 *tables*

Several authors (Kendall [7], Simpson [15]) warn against amalgamating 2^3 tables to 2×2 tables, even when second-order interaction happens to be absent. They demonstrate the possibility:

(a) that interactions, which in each of the amalgamated classes separately tend in the same direction, seem to be absent after amalgamation, or

(b) that the amalgamated classes separately do not show dependence between the two other properties, but that they do together. In our opinion this warning is exaggerated in many cases and danger threatens from another direction. In our view, as has been shown in Sec. 6, dependence between two classifications in a 2^3 table will just be tested in a 2×2 table obtained by amalgamation in the 2^3 table, irrespective of whether second-order interaction is present or not.

While second-order interaction is absent, the case (a) may be constructed by adding a level, a row $\times$ layer interaction, and a column $\times$ layer interaction. In the bottom of the table this appears as adding of a level and suitable row and column effects in a 2×2 table so that a small interaction occurs in it (see Sec. 3). In a similar way a small interaction which will have the same direction appears in the upper face. However, in the 2×2 table obtained by amalgamation of bottom and upper face, interaction will be absent.

Also, while second-order interaction is absent, the case (b) may be constructed in the following way. First we form the sum of a level with $p_3 = \frac{1}{2}$, a row $\times$ layer interaction, and a column $\times$ layer interaction. Because the interactions in the bottom and in the upper face of this sum are identical, we can add a row $\times$ column interaction to it such that both interactions in the bottom and in the upper face separately vanish. The 2×2 table obtained by amalgamation of bottom and upper face shows row $\times$ column interaction of course. These disturbing interactions, however, will be negligible in comparison with proper interactions due to dependence.

The danger to which we alluded consists in maintaining the hypothesis of independence in the model of a 2^3 table and the corresponding partition, although a first-order interaction appears to be considerable, or independence between one or more classifications can be expected

to be impossible in advance. For the discussed χ^2-test and the partition in a 2^3 table—and with this remark we proceed on what we said in Sec. 6—are only justified if the three classifications are independent.

In a case, e.g. like the example given by Simpson [15] where the result of a treatment against a disease is investigated by counting *dead* and *alive* in males and females, it is not obvious that the probability of being treated is the same for males and females.

If *one first-order interaction* must be taken in account, we have to choose a new model to test other interactions. The new model for this case is that for a 2×4 table. Let the experimental result multiplied by $n = \sum_{i=1}^{8} x_i$ be:

	treated		not treated	
	male	female	male	female
alive	x_1	x_2	x_5	x_6
dead	x_3	x_4	x_7	x_8

After the choice of a level similar to that in the foregoing section:

$$\begin{bmatrix} p_1p_2p_3p_4 & p_1p_2q_3p_4 & p_2q_2p_4 & q_1p_4 \\ p_1p_2p_3q_4 & p_1p_2q_3q_4 & p_1q_2q_4 & q_1q_4 \end{bmatrix},$$

the basis vectors for interactions may be:

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} p_3 & q_3 & -1 & 0 \\ -p_3 & -q_3 & 1 & 0 \end{bmatrix},$$

and

$$\begin{bmatrix} p_2p_3 & p_2q_3 & q_2 & 1 \\ -p_2p_3 & -p_2q_3 & -q_2 & -1 \end{bmatrix}.$$

For convenience we put the p_i as unknown for the rest of this section so that no main effects are present.

With the mentioned example in mind we prefer to define the vector level as:

$$\begin{bmatrix} p_1p_2p_4 & p_1q_2p_4 & q_1p_3p_4 & q_1q_3p_4 \\ p_1p_2q_4 & p_1q_2q_4 & q_1p_3q_4 & q_1q_3q_4 \end{bmatrix},$$

which is the same as that for a 2^3 table if $p_2 = p_3$. A basis for interactions may be formed by:

$$\begin{bmatrix} p_2 & q_2 & -p_3 & -q_3 \\ -p_2 & -q_2 & p_3 & q_3 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix},$$

and (15)

$$\begin{bmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

The first of these three vectors represents dependence of death rate from treatment, the second and the third represent interaction of sex and death rate within the treated and within the not-treated individuals respectively. If one prefers to consider dependence between sex and death rate and further dependence of death rate from the treatment within the sexes separately, one should only transpose the second and third column in each of the four vectors. We proceed in the former version.

The last pair of vectors can be replaced in several ways by two other vectors, the one representing a common interaction between sex and death rate in both treated and not-treated individuals, the other representing a difference in dependence of death rate from sex between treated and not-treated individuals, and, moreover, such that a partitioning of X^2 in independent components corresponds to this choice. We prefer:

$$\begin{bmatrix} p_1p_2q_2 & -p_1p_2q_2 & q_1p_3q_3 & -q_1p_3q_3 \\ -p_1p_2q_2 & p_1p_2q_2 & -q_1p_3q_3 & q_1p_3q_3 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}.$$

The first reason for this preference is that the mentioned difference in dependence—which may be called second-order interaction—is represented by the same basis vector as in the 2^3 table. A second reason will be given in the treatment of the following case. A third reason will appear later on. The component of X^2 corresponding to the first vector of (15) can be calculated as the test criterion in the 2×2 table obtained by neglecting sex. The computation of the other components will be treated in the following.

It is also possible that *two first-order interactions* must be taken in

account. Let us assume also that an interaction between treatment and death rate exists in the discussed 2×4 table. We take as a new level:

$$\begin{bmatrix} p_1p_2p_4 & p_1q_2p_4 & q_1p_3p_5 & q_1q_3p_5 \\ p_1p_2q_4 & p_1q_2q_4 & q_1p_3q_5 & q_1q_3q_5 \end{bmatrix}$$

which is the same as that for a 2×4 table if $p_4 = p_5$ and as that for a 2^3 table if, moreover, $p_2 = p_3$. This level corresponds to that for two separate and independent 2×2 tables. A basis for interactions may be formed by the last two vectors of the set given by (15).

This pair can be replaced again in several ways by another pair of vectors which express a common interaction between sex and death rate in both 2×2 tables, and a difference between such interactions respectively, and which admit a partition of X^2 in independent components. We choose:

$$\begin{bmatrix} p_1p_2q_2p_4q_4 & -p_1p_2q_2p_4q_4 & q_1p_3q_3p_5q_5 & -q_1p_3q_3p_5q_5 \\ -p_1p_2q_2p_4q_4 & p_1p_2q_2p_4q_4 & -q_1p_3q_3p_5q_5 & q_1p_3q_3p_5q_5 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}.$$

The difference in interaction is expressed again by the same vector as in the previous cases. The component of X^2 corresponding to the first vector of this pair is equal to:

$$\frac{(q_2q_4x_1 - p_2q_4x_2 - q_2p_4x_3 + p_2p_4x_4 + q_3q_5x_5 - p_3q_5x_6 - q_3p_5x_7 + p_3p_5x_8)^2}{n(p_1p_2q_2p_4q_4 + q_1p_3q_3p_5q_5)}.$$

According to the end of Sec. 5, this may be reduced to:

$$\frac{(x_1 - np_1p_2p_4 + x_5 - nq_1p_3p_5)^2}{np_1p_2q_2p_4q_4 + nq_1p_3q_3p_5q_5}.$$

As we know that a quantity like

$$\frac{x_1 - np_1p_2p_4}{\sqrt{np_1p_2q_2p_4q_4}}$$

(with n sufficiently large) has a standard normal distribution, we may consider x_1 as a normally distributed variable with expectation $np_1p_2p_4$ and variance $np_1p_2q_2p_4q_4$. The component is thus the square of a normally (0, 1) distributed combination of two normally (0, 1) dis-

tributed variables of the considered kind and such that the numbers np_1 and nq_1 have some weight in this combination, indeed, but not too much. This balanced combination of statistics for testing independence in 2×2 tables, which is also recommended by van Eeden [2], is the second reason for our preference mentioned in the previous case, which is obtained by equalizing p_5 to p_4 and q_5 to q_4 respectively. It may be remarked that the chosen measure of dependence agrees with Kendall's [7] quantities δ both in the case of a 2×4 table and in that of two 2×2 tables. The common interaction is thus not obtained by amalgamating the treated and not-treated classes, but by *adding* the two values of this measure of dependence, and this as a consequence of choosing the appropriate model.

The case where *three first-order interactions* must be taken in account will be considered now. If none of the faces of the original 2^3 table plays a special role in this case, we must have a definition of a level containing three interactions which is independent of the choice of the properties allotted to rows, columns, or layers respectively. After Bartlett [1] we choose a vector consisting of π_i $(i = 1 \cdots 8)$ with

$$\sum_{i=1}^{8} \pi_i = 1 \quad \text{and} \quad \pi_1\pi_4\pi_6\pi_7 = \pi_2\pi_3\pi_5\pi_8 = \gamma. \tag{16}$$

We remark that this relation was also valid for the levels of the 2^3 table, the 2×4 table, and the two 2×2 tables. Conceiving e.g. $\pi_1/\pi_2 : \pi_3/\pi_4 = \pi_1\pi_4/\pi_2\pi_3$ as a measure of interaction in the 2×2 table

$$\begin{bmatrix} \pi_1 & \pi_2 \\ \pi_3 & \pi_4 \end{bmatrix}$$

as is recommended for measuring linkage, we see that the relation (16) involves equality of interactions in every pair of faces of the original 2^3 table. The difference between the experimental vector and the maximum likelihood estimate of the level—which cannot be obtained in that simple way (i.e. from marginal totals) as in the cases discussed till now—will be called second-order interaction again. The corresponding X^2 has a one-dimensional χ^2-distribution because the estimation of the level implies the estimation of six parameters.

The estimates $\hat{\pi}_i$ are obtained by determination of the maximum of the likelihood function:

$$C + \sum_{i=1}^{8} x_i \log \pi_i \quad \text{under the conditions (16).}$$

Differentiation of the function:

$$\sum_{i=1}^{8} x_i \log \pi_i + \lambda(\pi_1\pi_4\pi_6\pi_7 - \pi_2\pi_3\pi_5\pi_8) - \mu\left(\sum_{i=1}^{8} \pi_i - 1\right)$$

with respect to π_i gives the equations:

$$P_i \equiv \frac{x_i + \lambda\gamma}{\pi_i} - \mu = 0, \qquad (i = 1, 4, 6, 7)$$

and

$$P_i \equiv \frac{x_i - \lambda\gamma}{\pi_i} - \mu = 0, \qquad (i = 2, 3, 5, 8).$$

$\sum_{i=1}^{8} \pi_i P_i$ yields: $\mu = n$. Putting $\lambda\gamma = \delta$, we obtain:

$$x_1 + \delta = n\hat{\pi}_1 , \qquad x_5 - \delta = n\hat{\pi}_5 ,$$

$$x_2 - \delta = n\hat{\pi}_2 , \qquad x_6 + \delta = n\hat{\pi}_6 ,$$

$$x_3 - \delta = n\hat{\pi}_3 , \qquad x_7 + \delta = n\hat{\pi}_7 ,$$

$$x_4 + \delta = n\hat{\pi}_4 , \qquad x_8 - \delta = n\hat{\pi}_8 .$$

From this it will be seen that second-order interaction is a multiple of the vector

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

here again, namely with coefficient δ/n. By substitution in (16) the following equation for δ is obtained:

$$(x_1 + \delta)(x_4 + \delta)(x_6 + \delta)(x_7 + \delta) = (x_2 - \delta)(x_3 - \delta)(x_5 - \delta)(x_8 - \delta).$$

Lancaster [12] showed that the test criterion $X^2 = (\delta^2/n) \sum_{i=1}^{8} 1/\hat{\pi}_i$ calculated after solution of this cubic equation is *asymptotically* equal to

$$\frac{\left(\frac{x_1}{\pi_1} - \frac{x_2}{\pi_2} - \frac{x_3}{\pi_3} + \frac{x_4}{\pi_4} - \frac{x_5}{\pi_5} + \frac{x_6}{\pi_6} + \frac{x_7}{\pi_7} - \frac{x_8}{\pi_8}\right)^2}{n \sum_{i=1}^{8} \frac{1}{\pi_i}}. \qquad (17)$$

Now the component of X^2 aimed at testing second-order interaction in the discussed models with no, one, and two interactions is equal to the same expression (17) for any n with the restriction only that the π_i may stand for estimates of the true π_i occurring in the level of the relative model which, however, converge to the true π_i for large n.

We see that the test criterion for second-order interaction belonging to any of the four discussed models—on condition that the underlying hypothesis expressed by the level is true—is asymptotically the same as that belonging to the following level if we observe the order of our treatment. This order implied that every model represented a stronger assumption about the π_i than the following models.

In other words, we may conclude that Bartlett's test of second-order interaction (admitting three first-order interactions) is asymptotically independent of whether any of these interactions is present or not; that our test of second-order interaction admitting two first-order interactions is asymptotically independent of whether any of these two interactions is present or not, but is not valid if three interactions occur in fact; that our test of second-order interaction admitting one first-order interaction is asymptotically independent of whether this interaction is present or not, but is not valid if one or two of the other true interactions is not zero; that the test of second-order interaction assuming no interactions (i.e. Lancaster's procedure) is only justified if no true interaction occurs in fact; that the value of any of the relative statistics is asymptotically equal to those admitting more interactions if only the null hypothesis belonging to the first statistic is not too narrow in the sense that interactions are supposed zero although they are present. This result, an extension of Lancaster's [12] remark, was the third reason for our preference in the choice of specific basis vectors.

Finally, a remark proceeding from a consideration of Bartlett's [1] numerical example also discussed by Lancaster [12]. This example showed a three-way classification of numbers of root-stocks according to time of planting (*at once* and *in spring*), to length of cutting (*long* and *short*), and to success (*alive* and *dead*):

	at once		in spring	
	long	short	long	short
alive	156	107	84	31
dead	84	133	156	209

Partition of the four-dimensional χ^2 corresponding to a 2^3 table yielded 95.58 for interaction between time of planting and success, 45.40 for interaction between length of cutting and success, 0.00 for interaction between length of cutting and time of planting, and 0.07 for second-order interaction. In connection with these large interactions, Bartlett's

and Lancaster's criteria are not equivalent and they will not be expected to be equal. If we formally follow the procedure discussed in this section, a new model, assuming two first-order interactions, would be needed for a further investigation of second-order interaction. We would consider these two 2×2 tables:

	alive		dead	
	long	short	long	short
at once	156	197	84	133
in spring	84	31	156	209

The two-dimensional χ^2, 7.41 (the sum of 6.50 and 0.91) could be partitioned in 5.26 for interaction between time of planting and length of cutting, and 2.15 for second-order interaction. Bartlett's criterion was equal to 2.27, so that the two criteria do not differ much now. The difference could be ascribed to a (formal) interaction between time of planting and length of cutting.

But in our opinion the whole procedure (also Bartlett's) seems to be wrong in this special example. For equality or unequality of dependence in the two considered 2×2 tables has no practical sense and will not be an object of investigation in this case. Moreover, the fact that the number of root-stocks is equal for all treatment combinations—which led up to an interaction χ^2 exactly equal to zero—suggests that these numbers were not random but fixed before the execution of the experiment. An interaction between time of planting and length of cutting must thus be excluded from the model. For that reason we have already referred to this interaction with the term *formal*.

In this and in similar cases we have to consider four independent binomial distributions defined by four chances π_i , in this example, chances of *alive* according to:

	at once		in spring	
	long	short	long	short
alive	π_1	π_2	π_3	π_4
dead	$1 - \pi_1$	$1 - \pi_2$	$1 - \pi_3$	$1 - \pi_4$

Here interactions are to be defined again. In the particular case where the numbers in every column are equal (as in this example), say n, the experimental result may be partitioned as follows:

$$\frac{1}{4n}\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} \pi & \pi & \pi & \pi \\ 1-\pi & 1-\pi & 1-\pi & 1-\pi \end{bmatrix}$$
$$+ \frac{\beta}{4}\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} + \frac{\gamma}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$
$$+ \frac{\delta}{4}\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}.$$

This partition corresponds to a partition of the three-dimensional test criterion for independence in a particular 2×4 table in three independent components. The second vector on the right represents interaction between time of planting and success, and the third vector, interaction between length of cutting and success. The sum of the first three vectors on the right is proportional to:

$$\begin{bmatrix} \pi+\beta+\gamma & \pi+\beta-\gamma & \pi-\beta+\gamma & \pi-\beta-\gamma \\ 1-\pi-\beta-\gamma & 1-\pi-\beta+\gamma & 1-\pi+\beta-\gamma & 1-\pi+\beta+\gamma \end{bmatrix},$$

i.e. a vector where $\pi_1 - \pi_2 = \pi_3 - \pi_4$, $\pi_1 - \pi_3 = \pi_2 - \pi_4$ and similar relations between $1 - \pi_i$ are valid. Such a vector where the (positive or negative) raising of the chance of *alive* or of *dead* by long cutting is the same for both times of planting, and where this raising by planting at once is the same for both lengths of cutting, seems to be a natural definition of the hypothesis *no second-order interaction* for this case. The fourth vector will represent *second-order interaction*, i.e. inequality of the mentioned raisings. The corresponding partition of X^2 is for this example numerically equivalent to that at the beginning of our remark about it. If the numbers in the columns are not equal, a first-order interaction, e.g. between time of planting and success, can be tested in the 2×2 table obtained by neglecting the other classification (length of cutting).

A test for second-order interaction will imply then (and also when the partition described shows considerable first-order interactions as in this example) a maximum likelihood estimation of the π_i under the hypothesis of "no second-order interaction," i.e. $\pi_1 + \pi_4 = \pi_2 + \pi_3$. To that end we determine the maximum of the likelihood function:

$$C + \sum_{i=1}^{4} x_i \log \pi_i + \sum_{i=1}^{4} (n_i - x_i) \log (1 - \pi_i) \quad \text{under that condition.}$$

Differentiating the function:

$$\sum_{i=1}^{4} x_i \log \pi_i + \sum_{i=1}^{4} (n_i - x_i) \log (1 - \pi_i) + \lambda(\pi_1 - \pi_2 - \pi_3 + \pi_4)$$

with respect to π_i yields four quadratic equations in π_1, π_2, π_3, and π_4 respectively. The usable solutions ($0 \leq \pi_i \leq 1$) are:

$$\hat{\pi}_1 = \frac{\lambda - n_1 + \sqrt{(\lambda - n_1)^2 + 4\lambda x_1}}{2\lambda},$$

$$\hat{\pi}_2 = \frac{\lambda + n_2 - \sqrt{(\lambda + n_2)^2 - 4\lambda x_2}}{2\lambda},$$

$$\hat{\pi}_3 = \frac{\lambda + n_3 - \sqrt{(\lambda + n_3)^2 - 4\lambda x_3}}{2\lambda},$$

$$\hat{\pi}_4 = \frac{\lambda - n_4 + \sqrt{(\lambda - n_4)^2 + 4\lambda x_4}}{2\lambda}.$$

Substitution in the relation between the π_i yields the following equation for λ:

$$\sqrt{(n_1 - \lambda)^2 + 4\lambda x_1} + \sqrt{(n_2 + \lambda)^2 - 4\lambda x_2} + \sqrt{(n_3 + \lambda)^2 - 4\lambda x_3} + \sqrt{(n_4 - \lambda)^2 + 4\lambda x_4} = \sum_{i=1}^{4} n_i$$

from which a solution different from the trivial solution $\lambda = 0$ is required, unless an approximate solution which may be given by

$$\frac{-\frac{x_1}{n_1} + \frac{x_2}{n_2} + \frac{x_3}{n_3} - \frac{x_4}{n_4}}{\sum_{i=1}^{4} \frac{x_i(n_i - x_i)}{n_i^3}}$$

and which can be improved by usual methods, is exactly zero. In that case the solution of λ is zero.

In Bartlett's example this approximate solution of λ was equal to 4.91. This could be improved to 4.92078. Solving the $\hat{\pi}_i$ with the help of this value gave the following table of expected numbers:

$$\begin{bmatrix} 157.11 & 105.78 & 82.89 & 31.56 \\ 82.89 & 134.22 & 157.11 & 208.44 \end{bmatrix}$$

The one-dimensional χ^2 for second-order interaction according to our definition of no second-order interaction appeared to be equal to 0.082.

ACKNOWLEDGEMENT

This work was begun while the author was assistant at the Department of Mathematics of the Agricultural University, Wageningen. He is much indebted to Professor Dr. N. H. Kuiper whose many suggestions and sound criticisms were an indispensable help to him.

REFERENCES

1. Bartlett, M. S. Contingency table interactions. *J. Roy. Stat. Soc. Suppl. 2*, [1935], 248.
2. Eeden, C. van. Methoden voor het vergelijken, toetsen en schatten van onbekende kansen. *Statistica 7*, [1953], 141–162.
3. Fisher, R. A. *Statistical Methods for Research Workers*, 12th ed., [1954], 101–106 and 299 et sqq.
4. Fog, D. Contingency tables and approximate χ^2-distributions. *Math. Scand. 1*, [1953], 93–103.
5. Haldane, J. B. S. Note on the preceding analysis of Mendelian segregations. *Biometrika 31*, [1939], 68.
6. Irwin, J. O. A note on the subdivision of χ^2 into components. *Biometrika 36*, [1949], 130–134.
7. Kendall, M. G. *The Advanced Theory of Statistics*, Vol. 1, [1945], 310–317.
8. Kimball, A. W. Short-cut formulas for the exact partition of χ^2 in contingency tables. *Biometrics 10*, [1954], 452–458.
9. Kuiper, N. H. Variantie-analyse. *Statistica 6*, [1952], 149–194.
10. Lancaster, H. O. The derivation and partition of χ^2 in certain discrete distributions. *Biometrika 36*, [1949], 117–129.
11. Lancaster, H. O. The exact partition of χ^2 and its application to the problem of the pooling of small expectations. *Biometrika 37*, [1950], 267–270.
12. Lancaster, H. O. Complex contingency tables treated by the partition of χ^2. *J. Roy. Stat. Soc. B 13*, [1951], 242–249.
13. Lancaster, H. O. A reconciliation of χ^2, considered from metrical and enumerative aspects. *Sankhya 13*, [1953], 1–10.
14. Roberts, E., Dawson, W. M., and Madden, M. Observed and theoretical ratios in Mendelian inheritance. *Biometrika 31*, [1939], 62.
15. Simpson, E. H. The interpretation of interaction in contingency tables. *J. Roy. Stat. Soc. B 13*, [1951], 238–241.

THE DISTRIBUTION OF RED BLOOD CELLS IN THE HEMACYTOMETER

Malcolm E. Turner

Medical College, University of Cincinnati, Cincinnati, Ohio, U. S. A. and Institute of Statistics, North Carolina State College, Raleigh, North Carolina, U. S. A.

G. S. Eadie

Duke University School of Medicine, Durham, North Carolina, U. S. A.

1. *Introduction*

The validity of the hemacytometer method for estimating the density of erythrocytes depends upon several factors. Among these, it is necessary to assume that the aliquot of blood taken from the organism is truly representative of the total blood mass. In addition, it must be assumed that the erythrocytes in the aliquot become dispersed completely at random over the surface of the counting chamber. The manner in which the erythrocytes randomly distribute themselves into the various square divisions of the chamber grid becomes of interest by virtue of the fact that knowledge of this distribution can yield information concerning the precision of the estimated number of cells. Abbé [1878] seems to be the first to suggest that there is a relation between precision and the magnitude of the count itself. He proposed using as a measure of precision the square root of the mean number of cells per division as a consequence of the distribution law which he derived. This distribution law was in fact identical to the well-known Poisson distribution [1837]. Student [1907] independently derived Poisson's distribution in regard to counting yeast cells.

It is well known that the mean number of cells per division is a good estimate, in fact the maximum likelihood estimate, of the variance of a Poisson distribution; so that the standard deviation might simply be taken as the square root of the mean as Abbe proposed. However, in more recent years, Berkson [1938] has calculated the mean and variance of a number of cases of actual red blood cell counts made under very excellent counting conditions and has shown that the variance is almost invariably smaller than the mean. Berkson tested the equality

of the mean and variance, implied by the Poisson distribution, on this data by the use of Fisher's [1922] chi-square test.

Lancaster [1952] investigated the use of this chi-square criterion for controlling counting and found it quite adequate to the task. Berkson and his associates discuss the adequacy of the Poisson model in a number of papers [1935, 1936, 1938, 1939, and 1940]. Whitaker [1914] and Student himself [1919] both give illustrations of the inadequacy of the Poisson in some particular instances and show that in many of the common applications of the distribution the underlying assumptions are not even approximately met.

Lancaster [1950] attacked the problem of crowding of the erythrocytes which had been noted by Berkson, *et al.* [1935]. He derived the consequences of crowding upon the relationship of the variance to the mean but did not offer a complete distribution law to substitute for the Poisson distribution.

Plum [1936] should be consulted for the early history of the subject and Lancaster for later developments [1950].

2. *The Poisson-Binomial Model*

It is a well known fact that the Poisson distribution can be considered to be a limiting form of the binomial distribution. This latter distribution can be derived for the hemacytometer situation in the following manner.

Let us suppose that there are a total of r blood cells which finally come to rest under n squares of the hemacytometer grid. We would like to know in what manner these r cells will distribute themselves into the squares of the grid. In other words, what is the probability that there will be no cells in a given square; what is the probability that there will be one cell in the given square; or two cells in the given square; or three cells; or in general k cells in the given square? Now if we assume that all of the r cells are distinguishable and that all n^r possible arrangements are equally likely to occur, then the probability of obtaining exactly k cells in a given square can be found.

It is seen that this probability $P(k)$ is simply the

$$\frac{\text{No. of arrangements providing } k \text{ cells in the given square}}{\text{Total no. of possible arrangements}}$$

Now, there are $\binom{r}{k}$ ways of selecting k cells out of r cells and for each of these ways there is $(n-1)^{r-k}$ ways of arranging the remaining

$r - k$ cells into the remaining $n - 1$ squares. Thus,

$$P(k) = \frac{\binom{r}{k}(n-1)^{r-k}}{n^r}. \tag{1}$$

This is the general term of the series for the binomial $[1/n + (n-1)/n]^r$.

In the hemacytometer situation $n = 400$ if we confine our attention to the smallest subdivisions in one large square of the standard hemacytometer. Each of these smallest squares has an area of 1/400th of a square mm. The total number of cells under the grid, r, is usually of the order of one to three thousand. With values of n and r in this range it can be easily demonstrated that the general term of the Poisson distribution

$$\frac{1}{k!}\exp\left\{-\frac{r}{n}\right\}\left(\frac{r}{n}\right)^k \tag{2}$$

can be used as a quite good approximation for $P(k)$. Now it is well known that the variance of the binomial distribution is rpq which in the notation of this paper becomes

$$r\left(\frac{1}{n}\right)\left(1 - \frac{1}{n}\right); \tag{3}$$

but this is, to a close approximation when $n = 400$ just r/n, which is, in fact, the mean. It is also true, as previously mentioned, that the variance of the Poisson distribution is equal to the mean. Thus, it can be seen that if this binomial-Poisson model were appropriate we should expect the estimated mean and the estimated variance to be of equal magnitude on the average. This is not the case, as Berkson showed. The difficulty can be traced to the assumptions utilized in deriving the model. We assumed that all n^r possible arrangements are equally likely. This implies that there is no crowding; that is, if a cell occupies a square, that it is just as likely that another cell would come to rest in that square as come to rest in any other square, one, say, that has not been previously occupied. Now, this is a reasonable assumption so long as a square is not nearly filled with cells. However, with the range of r usually encountered in red cell hemacytometry and erythrocytes being as big as they are, it happens that a number of squares get quickly filled and this results in a more even dispersion of cells throughout the counting chamber. This, obviously, results in a lessened observed variance than would be predicted by the binomial-Poisson model.

3. *An Improved Model*

We will discuss in this paper a model which takes into account the above-mentioned crowding due to the finite limitations of space in the counting chamber.* A derivation can be made as follows.

Let us suppose that N is the maximum number of cells that can be accommodated under the entire grid, assuming no over-lapping of cells. We denote the maximum number of cells which can be accommodated under one square of the grid as a. Obviously $a = N/n$ since each of the n squares has equal area. We desire the probability, $P(k_1 , k_2 , \cdots , k_n)$ of obtaining exactly $k_1 , k_2 , \cdots , k_n$ cells in the n squares of the grid.

Now, the total number of ways in which the r cells can be arranged in the N available spaces is $\binom{N}{r}$, since each space can contain only one cell. We assume that each of these ways has an equal chance of occurring. The number of ways that k_1 cells can be arranged in the a spaces available in the first square, k_2 cells in the second, and so on is

$$\binom{a}{k_1}\binom{a}{k_2} \cdots \binom{a}{k_n}.$$

Then, the desired probability, of obtaining exactly $k_1 , k_2 , \cdots , k_n$ cells in the n squares is given by

$$\frac{\binom{a}{k_1}\binom{a}{k_2} \cdots \binom{a}{k_n}}{\binom{N}{r}} \tag{4}$$

where the sum of the k's is equal to r.

It is tacitly implied by this derivation that all of the r cells are of the same size. Of course, erythrocytes, like all biological products, vary somewhat in size; however, it seems likely that the model (4) will still serve as a reasonably good approximate model. The adequacy of the approximation, as in all scientific models, must be determined by the ability of the model to describe approximately all relevant details of the observational record.

---

*We are obligated to Dr. E. J. Williams for suggesting the model discussed here. The authors had employed the hypergeometric distribution given by

$$P(k) = \binom{a}{k}\binom{N-a}{r-k} \Big/ \binom{N}{r}$$

in an earlier draft of this paper. This latter was not so general in that the correlation between squares, which had been deemed negligible, was not taken into account. Formula (5) below was also provided by Dr. Williams.

It can be shown that the expected value of the estimated variance of k_i $(i = 1, 2, \cdots, n)$ is given by

$$\sigma^2 = \frac{N}{(N-1)}\mu - \frac{n}{N-1}\mu^2, \tag{5}$$

where μ is the expected mean, r/n. When N is large, the coefficient of μ in the first term of formula (5) becomes close to unity so that the variance can be approximately expressed by

$$\sigma^2 \approx \mu - \frac{n}{N-1}\mu^2. \tag{6}$$

This is to be compared with the corresponding expression for the Poisson distribution,

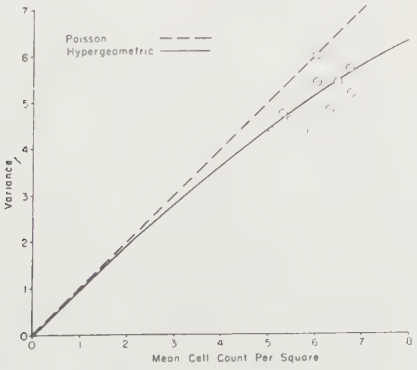
$$\sigma^2 = \mu \tag{7}$$

Formula (5) should be compared with the quadratic relationship derived by Lancaster (1950) by a non-distributional argument. In his approach to the crowding problem Lancaster was working with larger subdivisions of the hemacytometer grid than we have employed and his derivation was geometrical in nature. It is interesting that these two independent developments have produced theoretical relationships between the mean and variance of the same algebraic form.

4. *Test of the Model*

Formula (6) provides a convenient way of testing the adequacy of the suggested model. We do not know the value of N but the constant $n/(N-1)$ can be estimated from actual data by fitting the quadratic equation (6). Figure 1 shows the data of Berkson [1935] with the best

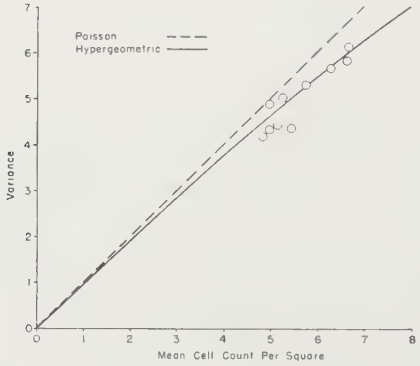
FIGURE 1
RELATIONSHIP BETWEEN THE VARIANCE AND THE MEAN CELL COUNT (DATA OF BERKSON)



fitting (in the sense of least squares) quadratic curve drawn through the points. The straight line drawn at 45° from the axes represents the relationship between the mean and variance expected on the basis of the Poisson model. The description of the data seems to be quite good.

A similar series of observations was made at Duke University under conditions rivaling those of Berkson in accuracy of counting. These data are presented in Figure 2, along with the Poisson line and

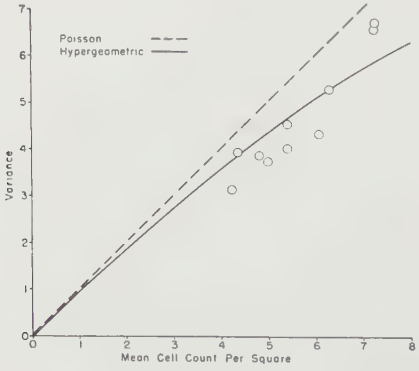
FIGURE 2
RELATIONSHIP BETWEEN THE VARIANCE AND THE MEAN CELL COUNT (DUKE SERIES #1)



the quadratic relationship for the more general model. Again, the description appears to be not at all bad.

Both the series of Berkson and the Duke series represent observations on samples of blood from different individuals. In Figure 3

FIGURE 3
RELATIONSHIP BETWEEN THE VARIANCE AND THE MEAN CELL COUNT (DUKE SERIES # 2)



is presented a series of observations made at Duke collected in a different manner. Each mean and variance is based on counts of erythrocytes for the same blood sample. The mean number of cells per square has been varied by using various degrees of dilution of the single sample of blood. The dilutions used were 1:250, 1:225, 1:200, 1:175, and 1:150. Two preparations were made at each of these dilutions. Each of the ten resulting means and variances is based upon 208 of the small squares instead of 400 as in the two previously described series. The description by the quadratic curve would be quite good if it were not for the two high values. Perhaps these two values represent just random variation, but we are inclined to think that their departure from expectation may be due to disturbances in the chamber filling process. As we intend to show elsewhere, the variance of the hemacytometer count can be increased to a considerable extent by interruption in the filling of the counting chamber so that a homogeneous distribution of cells in the chamber does not result.

5. *Size of the Erythrocytes*

We have attempted to test the tenability of the new model by comparing observation with expectation as regards one particular consequence of this model, namely, the expected quadratic relationship between the variance and the mean. We propose the conclusion that the results are not, on the whole, at all inconsistent with the model. It certainly appears that the new model provides a better description of the presented data than the Poisson model. There is another consequence of the suggested model which can easily be put to a test. We have mentioned the fact that it is the effect of crowding which diminishes the variability of the number of cells per square. Now, crowding can be produced in two different ways. We can either increase the number of cells which compete for the available spaces or we can increase the size of the cells. The data previously discussed demonstrate the effect of crowding by increasing the number of cells. We now consider the alternative situation, the effect of cell size.

First, we note that it is possible to estimate N directly by use of the following relationship which is derived from formula (5) by rearranging terms:

$$N = \frac{n\mu^2 - \sigma^2}{\mu - \sigma^2} \tag{8}$$

where we substitute the observed mean and variance for σ^2 and μ. This is, of course, the "moment" estimate and is, perhaps, not the most efficient estimate of N. However, since the maximum likelihood

estimator is considerably more difficult to apply, we will use this simple estimator to illustrate our point. Now, N is by definition the total number of available spaces, so that it is related inversely to the average cell size. If we know that we have a sample of blood containing large cells we should expect our estimate of N to be smaller than that for normal sized cells. The reverse, of course, should hold true if we have a sample containing smaller than average cells. Failure of the results to comply to this expectation would throw some suspicion on the adequacy of the model.

A sample of blood was obtained from a patient with microcytic anemia and from one with macrocytic anemia. The number of cells per square was counted for each of the 400 small squares and the mean and variance were calculated. These statistics were obtained for both patients. N, the number of available spaces, was estimated by formula (8) and the results are given in Table 1. The average cell diameter

TABLE 1

Patient	No. of squares counted	Mean No. of cells	Variance	$\hat{N}$	Estimated cell diameter (in microns)	Mean corpuscular volume
Microcytic	400	5.548	5.045	25,100	6.3	65
Macrocytic	400	3.420	3.242	14,900	8.2	104

can be approximated from N (and hence from the mean and variance) by assuming that the cells each occupy a square with sides equal to the diameter of the cell. Thus, the cell diameter in microns is approximaterly given by

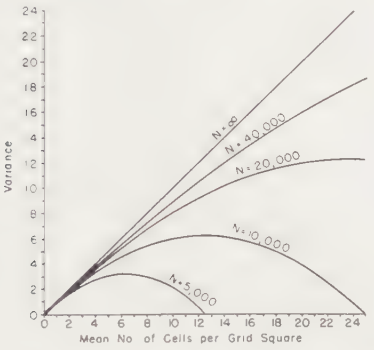
$$\text{Cell diameter} \approx 1000/\sqrt{N}. \tag{9}$$

The mean corpuscular volume has been computed from the hematocrit and cell count for comparative purposes. We are not proposing calculation of the cell diameter from the mean and variance as a practical procedure; but it appears to us to yield an interesting confirmation of the essential validity of the model. The actual average cell diameters are not known but the values computed above are of a reasonable magnitude for patients of these two types.

6. *Discussion*

The combined effect on the variance of crowding by increasing the size of the erythrocytes, which is equivalent to decreasing N, and of increasing the density of erythrocytes in the counting chamber, may be

FIGURE 4
RELATIONSHIP BETWEEN THE VARIANCE AND MEAN CELL COUNT FOR A VARIETY OF HYPOTHETICAL CASES



seen from Figure 4. The range of values, both as to numbers and size, is considerably larger than is encountered in actual practice. However, this extension of the range makes possible a portrayal of the full implications of the suggested model. For example, it is apparent that as the density of cells is increased and/or the size of the cells is increased, the variance of the count increases to a point, and then with further increase of density or size, the variance declines and eventually becomes zero when all of the chamber squares are filled to a maximum and hence all contain the same number of erythrocytes. The line corresponding to $N = \infty$ is the degenerate case, the Poisson distribution. When N is infinite, then the cells must occupy no spaces at all, and hence no crowding occurs. Thus, by considering the Poisson as a limiting case of the suggested distribution, we realize why this simple distribution does not suffice to describe the distribution of such relatively large objects as erythrocytes when in considerable density.

The reader may wonder why we have chosen to test the theory by such indirect measures as those we have described. Why not simply test the goodness of fit of the theoretical distribution with the observed distribution? This was actually done, but it turned out to be a quite insensitive mode of discrimination. It happens that several different distributions, including both the Poisson and the suggested distribution, yield fits to the data which are all quite good by the chi-square goodness of fit criterion. However, it is not the overall fit, but rather the precision of the count with which we are primarily concerned. It is this specific aspect of the description wherein the new model appears to be an improvement over the Poisson.

We will terminate this discussion by a few comments as to the practical implications of the above theory. We were motivated in this

investigation by the desirability of being able, in actual practice, to place confidence limits on cell counts without having actually to bother with making a separate estimate of the variance. By assuming that the cells were distributed in the counting chamber according to the Poisson Law, we were able to do just that. In this case the variance is taken to be the count itself. Now, if instead we assume that the law (4) holds, then it is not possible to know the variance from the count alone. In addition, we must know the size of the erythrocytes; however, this information is in general not known. Fortunately, there is a very simple solution to this dilemma, one that is apparent from the graphs of the relationship between the variance and mean. If the cell suspension is sufficiently diluted, then the limiting Poisson distribution becomes an adequate approximation to the true distribution and the variance of the count can be taken to be the count itself.

Another solution might be considered, but turns out to be of no value for our purposes. This solution would be to increase the density of cells to the point that the variance approaches zero. Certainly, we should desire the variance to be as small as possible. However, the effect of this procedure is to destroy the representativeness of the sample in that there would be a piling-up of cells and thus our necessary condition of randomness would be lost.

7. *Summary*

Thus, we finally arrive back where we started via a rather devious route. We began with the empirical evidence against the use of the Poisson distribution. Then we derived a new model with somewhat less unrealistic assumptions about the physical system we were studying. We saw that this latter model fairly well represented a variety of actual data. Then we considered from an intuitive point of view the known fact that the Poisson is a limiting form of this latter distribution. We then concluded that the Poisson can be made to apply to red cell hemacytometry by making greater dilutions of the cells than is at present generally done. This conclusion is justified in two different ways. First, with sufficient dilution we can use the Poisson distribution; and, second, it seems that there is less chance of disturbing the randomness of the sample by eliminating a tendency to pile-up.

ACKNOWLEDGEMENTS

The authors wish to thank Ivan W. Brown, Jr. for making available the resources of his laboratory; R. L. Anderson, H. O. Lancaster, H. L. Lucas, R. J. Monroe, S. N. Roy and E. J. Williams for reviewing various portions of the argument; and W. Grandy Curtis for his technical

assistance. Especial thanks are due Dr. E. J. Williams for his suggested modification of the authors' proposed model.

This investigation was supported in part by a research grant [H1226(C5)] from the National Heart Institute of the National Institutes of Health, Public Health Service.

REFERENCES

Abbé, E. [1878]. Ueber Blutkörper-Zahlung. *S. B. Jena. Ges. Med. Naturw. no. 48.*

Berkson, J. [1938]. Some difficulties of interpretation encountered in the application of the chi-square test. *J. Amer. Stat. Assoc. 33*: 526.

Berkson, J. [1940]. A note on the chi-square test, the Poisson and the binomial. *J. Amer. Stat. Assoc. 35*: 362.

Berkson, J., Magath, T. B. and Hurn, M. [1935]. Laboratory standards in relation to chance fluctuations of the erythrocyte count as estimated with the hemacytometer. *J. Amer. Stat. Assoc.* 30: 414.

Berkson, J., Magath, T. B. and Hurn, M. [1939]. The error of estimate of the blood cell count as made with the hemacytometer. *Amer. J. Physiol. 128*: 309.

Fisher, R. A., Thornton, H. G. and Mackenzie, W. A. [1922]. The accuracy of the plating method of estimating the density of bacterial populations. *Ann. Appl. Biol. 9*: 325.

Lancaster, H. O. [1950]. Statistical control in haematology. *Journal of Hygiene 48*: 402.

Lancaster, H. O. [1952]. Statistical control of counting experiments. *Biometrika 39*: 419.

Magath, T. B., Berkson, J. and Hurn, M. [1936]. The error in determination of the erythrocyte count. *Amer. J. Clin. Path. 6*: 568.

Plum, P. [1936]. Accuracy of haematological counting methods. *Acta. Med. Scand. 90*: 342.

Poisson, S. D. [1837]. Recherches sur la probabilité des jugements. Paris, 205.

Student [1907]. On the error of counting with a haemocytometer. *Biometrika 5*: 351.

Student [1919]. An explanation of deviations from Poisson's law in practice. *Biometrika 12*: 211.

Whitaker, L. [1914]. On Poisson's law of small numbers. *Biometrika 36*: 36.

THE COMPARISON OF THE SENSITIVITIES OF SIMILAR EXPERIMENTS: APPLICATIONS*

R. A. Bradley and D. E. W. Schumann**

Virginia Agricultural Experiment Station of the Virginia Polytechnic Institute, Blacksburg, Virginia, U. S. A.

INTRODUCTION

The problem of deciding on the best method of measuring the effects of experimental treatments is a difficult one in certain areas of research. We are concerned with means of comparing the sensitivities of two available experimental techniques, where we use the term in a very general sense, that might be used to yield data for specified treatment comparisons. It is often not possible to determine on a priori grounds which of two techniques should be used and it may be desirable to use both techniques in separate experiments with a view to deciding on the technique to be used in future experimentation. When the two experiments are indeed independent of each other in probability, the methods of this paper may be used.

Let us consider several typical situations indicated by subject-matter headings.

Taste Testing: Hopkins [1950, 1953] investigated the use of discrete scoring scales in taste testing experiments as they are commonly used and Baten [1946] suggested the use of a continuous line scoring system. The latter obtained results which he interpreted as showing an increase in sensitivity for his method over the use of discrete scores. Procedures discussed in this paper provide means of comparing the sensitivities of the two scoring systems if data were available from two identically designed but statistically independent experiments, one for each scoring system, in a form suitable for analysis of variance.

Many other comparable situations arise in taste testing. The sensitivities of two taste panels, two methods of preparation of test items, two environmental situations during tasting, or two methods of storage may be compared.

*Research sponsored by the Agricultural Research Service, U.S.D.A. under a Research and Marketing Act Contract, No. 12-14-100-126(20).

**Now with the Department of Statistics, University of Stellenbosch, Stellenbosch, South Africa.

Chemistry: Consider the analytical chemist with two series of reactions based on the use of different reagents available for a quantitative analysis of a particular element in a compound. His ultimate objective is to run comparative experiments on the concentrations of this element in samples of the compound obtained from a number of suppliers. Two similar experiments might be conducted, one for each analytical method, to determine which method is to be used in the required routine experiments. Other problems of this nature arise frequently in chemistry and all of the physical sciences.

Life Testing: Means of conducting accelerated life tests are being devised in electronics, food processing, nuclear physics, and in a host of other research endeavors. In some cases it may be possible to compare experiments based on accelerated tests with experiments depending on normal life tests. In other cases alternative means of conducting accelerated tests may be available and comparisons of their sensitivities in the sense of emphasizing real treatment differences are required.

Agronomy: One of the difficult areas of agronomic research is in pasture experimentation. Yields of various forage mixtures may be measured through clipping of plots in some experiments while in others animal performance or carrying capacity may be taken as the measure of yield. While care should be taken to determine that these yield measurements may be transformed to comparable units, the need for comparison of the sensitivities of techniques relative to the exhibition of differences between measurable forage characteristics is again apparent.

Cochran [1943] discussed the problem of comparing different scales of measurement for experimental results. He assumed that analysis of variance techniques were applicable and confined his attention to the case in which all scales measure the same experiment. It was noted that a comparison of the sensitivities of two scales should depend both on the experimental errors associated with them and on the magnitudes of the treatment effects in the scales. Cochran suggested that a comparison of this sort should depend on a test of significance of the hypothesis that the parameters of two noncentral variance-ratio distributions are equal. Schumann and Bradley [1957] have provided means of making such tests of significance for comparisons of the sensitivities of two identically designed independent experiments. The main results are then for identical experiments in the sense that both experiments yield F-tests with the same degrees of freedom and with treatment means based on the same number of replications although generalizations are indicated. We shall summarize that work in the next section and devote the rest of this paper to numerical illustrations.

We may have somewhat generalized Cochran's concept of comparing scales of measurement in discussing comparisons of experimental techniques. Now some caution is in order in interpreting the results on comparisons of experiments. The test procedures developed are correct whenever two independent variance ratios are compared. However, the experimental techniques under comparison should be such that they do not interact with the treatments under comparison in a way not explainable in terms of a shift in treatment locations and a magnification of differences in treatment effects. It would be a good rule to forego experiment sensitivity comparisons if there is evidence of treatment by experiment interaction of a nature other than that permitted in view of our preceding statement.

Many experiments are conducted with a view to examining interactions. To illustrate, suppose that two independent experiments in agriculture differ only in geographical location. Then it is more likely that the experimenter will be interested in treatment by location interactions than in location sensitivity differences to treatment effects. The important applications of our methods seem to be in comparing experiments based on experimental techniques differing essentially in means of measurement which do not interact fundamentally with the treatments under comparison. It is of course obvious, if two experimental techniques are to be compared, that the comparison should not be confounded with other experimental changes.

The method for the comparison of the sensitivities of experiments may also be used to compare two population multiple correlation coefficients whose sampling distributions are based on regression models. An example of this method will be given.

DISTRIBUTION THEORY

The following summary of the distribution theory for the comparison of the sensitivities of two experiments is based on the work of Schumann [1956] and Schumann and Bradley [1957].

Let F_1 and F_2 be independent noncentral variance ratios, each with $2a$ and $2b$ degrees of freedom and with parameters of noncentrality, λ_1 and λ_2. Then their distributions are of the form

$$f(F_i) = (a/b)^a[B(a, b)]^{-1}e^{-\lambda_i}F_i^{a-1}(1 + aF_i/b)^{-(a+b)} \cdot {}_1F_1[a + b, a, a\lambda_iF_i/b(1 + aF_i/b)], \quad 0 \leq F_i \leq \infty, \; i = 1, 2, \tag{1}$$

where B represents the beta function and ${}_1F_1$, the confluent hypergeometric series. In Model I of the analysis of variance with so-called "fixed" parameters in the additive model, if τ_{ij} is the effect of the j-th

of t treatments in experiment i and if σ_i^2 is the i-th population experimental error variance,

$$\lambda_i = k \sum_{j=1}^{t} \tau_{ij}^2/2\sigma_i^2 , \qquad i = 1, 2, \tag{2}$$

when k is the number of observations in each treatment mean. These parameters of noncentrality are then measures dependent upon both the experimental error associated with the experiment and the magnitudes of treatment effects in the experiment. When $\lambda_i = 0$, $f(F_i)$ in (1) becomes the density function of a central variance ratio with $2a$ and $2b$ degrees of freedom. The specification of degrees of freedom as $2a$ (for the numerator of F_i) and $2b$ (for the denominator of F_i) was for mathematical convenience; a and b may be integers or half-integers.

We required the distribution of

$$w = F_1/F_2 \tag{3}$$

and its density function is

$$g(w; a, b, \lambda) = e^{-2\lambda} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\lambda^{r+s}}{r!s!} [B(a + r, b)B(a + s, b)]^{-1} \cdot w^{a+r-1} H(w; r, s), \qquad 0 \leq w \leq \infty , \tag{4}$$

when

$$\lambda_1 = \lambda_2 = \lambda \tag{5}$$

and where

$$\begin{aligned} H(w; r, s) &= \int_0^{\infty} y^{2a+r+s-1}(1 + wy)^{-(a+b+r)}(1 + y)^{-(a+b+s)}\, dy, \qquad 0 \leq w \leq \infty , \\ &= B(2a + r + s, 2b)\,{}_2F_1(a + b + r, 2a + r + s, \\ &\qquad 2a + 2b + r + s, 1 - w), \qquad 0 \leq w \leq 1, \\ &= w^{-(2a+r+s)} B(2a + r + s, 2b) \\ &\cdot\, {}_2F_1[a + b + s, 2a + r + s, 2a + 2b + r + s, (w - 1)/w], \\ &\qquad 1 \leq w \leq \infty . \end{aligned} \tag{6}$$

${}_2F_1$ in (6) is the hypergeometric function. It has been noted that $g(w; a, b, \lambda)$ has a mode at a value of w between 0 and 1, a median at

TABLE 1

VALUES OF w_0 FOR SIMILAR EXPERIMENTS SUCH THAT $1 - G(w_0) = .05$

(Republished from Table I, Schumann and Bradley [1957], with the permission of the Editor, Annals of Mathematical Statistics)

	b	1	2	3	4	5	6	7	8	9	10	15	∞
a	d.f.	2	4	6	8	10	12	14	16	18	20	30	∞
1	2	66.12	32.76	26.76	24.37	23.10	22.31	21.77	21.39	21.10	20.87	20.21	19.00
3/2	3	40.81	23.15	14.40	12.81	11.97	11.45	11.09	10.85	10.65	10.50	10.07	9.28
2	4		13.91	10.62	9.32	8.62	8.19	7.90	7.69	7.54	7.41	7.06	6.39
5/2	5		11.82	8.87	7.70	7.11	6.68	6.42	6.23	6.09	5.96	5.64	5.05
3	6			7.86	6.77	6.18	5.78	5.54	5.37	5.25	5.13	4.82	4.28
7/2	7			7.22	6.17	5.61	5.26	5.03	4.85	4.72	4.61	4.30	3.79
4	8				5.75	5.21	4.88	4.65	4.48	4.35	4.24	3.94	3.44
9/2	9				5.45	4.92	4.59	4.36	4.19	4.07	3.97	3.68	3.18
5	10					4.70	4.37	4.14	3.98	3.86	3.76	3.48	2.98
11/2	11					4.52	4.19	3.98	3.82	3.70	3.60	3.32	2.82
6	12						4.05	3.84	3.68	3.56	3.46	3.18	2.69
13/2	13						3.94	3.72	3.57	3.45	3.35	3.07	2.58
7	14							3.63	3.47	3.36	3.26	2.98	2.49
15/2	15							3.54	3.39	3.27	3.18	2.90	2.40
8	16								3.32	3.20	3.11	2.83	2.34
17/2	17								3.26	3.14	3.05	2.77	2.28
9	18									3.09	3.00	2.72	2.22
19/2	19									3.04	2.95	2.67	2.17
10	20										2.91	2.63	2.12
21/2	21										2.87	—	—
15	30											2.36	1.85

$w = 1$ and that $\lim g(w) = 0$ as $w \to 0$ and as $w \to \infty$. Further, we use the fact that

$$g(w; a, b, \lambda) = g(1/w; a, b, \lambda). \tag{7}$$

We have shown that the distribution function

$$G(w_0 ; a, b, \lambda) = P(w \leq w_0 | a, b, \lambda) = \int_0^{w_0} g(w; a, b, \lambda)\, dw$$

may be approximated by the distribution function,

$$G(w_0; a', b, 0) = [B(a', b)]^{-2} \int_0^{\infty} y^{a'-1}(1 + y)^{-(a'+b)}(w_0 y)^{a'}(1 + w_0 y)^{-a'} \cdot \left[\frac{1}{a'} - \frac{(b-1)w_0 y}{(a'+1)(1+w_0 y)} + \frac{(b-1)(b-2)(w_0 y)^2}{2!(a'+2)(1+w_0 y)^2} - \cdots\right] dy \tag{8}$$

where

$$a' = (a + \lambda)^2/(a + 2\lambda). \tag{9}$$

The suitability of this approximation was determined on the basis of comparing moments of $g(w; a, b, \lambda)$ and of $g(w; a', b, 0)$ and a limited number of values of $G(w_0 ; a, b, \lambda)$ and $G(w_0 ; a', b, 0)$.

A table of values of $w_0(.05)$ for which $G[w_0(.05); a, b, 0] = .95$ was prepared by Schumann and Bradley [1957, Table I] for integer and half-integer values of a, $a = 1, \cdots, 21/2, 15$ and for integer values of b, $b = 1, \cdots, 10, 15, \infty$. This table is reproduced here as Table 1 for the convenience of the reader. The table is essentially triangular for $G(w_0 ; a, b, 0) = G(w_0 ; b, a, 0)$. It is apparent from the table that values of $G(w_0 ; a, b, \lambda)$ and $G(w_0 ; a', b, 0)$ are fairly stable even for considerable variation in values of λ and for quite small values of a and b. This implies that it will have little effect in applications if we enter the table using a value of λ somewhat different from its true (and usually unknown) value. Linear interpolation in Table 1 will be satisfactory in applications.

In usual multiple regression theory with non-stochastic independent variables, if we fit two regression equations of the form

$$y = a + b_1 x_1 + \cdots + b_p x_p \tag{10}$$

to two independent sets of data and obtain corresponding multiple correlation coefficients, R_1 and R_2 , based on N observation vectors each,

$$F_i = bR_i^2/a(1 - R_i^2), \qquad i = 1, 2 \tag{11}$$

have distributions of the form (1) with

$$\lambda_i = (a + b)\rho_i^2/(1 - \rho_i^2)$$
$$a = p/2 \tag{12}$$
$$b = (N - p - 1)/2$$

where ρ_i , $i = 1, 2$, is the population multiple correlation coefficient corresponding to R_i . It then follows that the procedures developed apply here and that we may perform tests of hypotheses on multiple correlation coefficients dependent on those procedures. It is common practice to use R^2 as a measure of the fraction of the variability in the dependent variable explained by regression. It will be useful in many cases to be able to test hypotheses on the equality of ρ_1^2 and ρ_2^2 .

Note that the application to comparing multiple correlation coefficients is suitable for the regression model (10) but does not apply to correlations based on models assuming a multivariate normal distribution for vectors $(y, x_1 , \cdots , x_p)$, for then the distribution (1) is not correct.

TEST PROCEDURES

(*i*) *Tests of hypotheses on sensitivity*

In tests of sensitivity we shall use the statistic $w = F_1/F_2$ defined in (3) and on the assumptions stated earlier that F_1 and F_2 are independent, have $2a$ and $2b$ degrees of freedom, and have noncentral F-distributions with parameters λ_1 and λ_2 . $w_0(\alpha)$ is defined to be such that $G[w_0(\alpha); a, b, \lambda] = 1 - \alpha$. Equality of λ_1 and λ_2 implies equal sensitivities for the two experiments so long as k in (2) is the same for each experiment, that is, so long as treatment means are based on equal numbers of replications.

$$\text{Test 1.} \qquad H_0 : \lambda_1 = \lambda_2 = \lambda, \qquad H_{a_1} : \lambda_1 > \lambda_2 . \tag{13}$$

The null hypothesis H_0 represents the hypothesis of equal sensitivities for the two identical experiments. The alternative hypothesis H_{a_1} assumes that we are interested in the one-sided test perhaps because we shall not change to the method of Experiment 1 from the currently used method of Experiment 2 unless the former is more sensitive. Given α, we obtain $w_0(\alpha)$ and reject H_0 in favor of H_{a_1} at significance level α when $w \geq w_0(\alpha)$.

$$\text{Test 2.} \qquad H_0 : \lambda_1 = \lambda_2 = \lambda, \qquad H_{a_2} : \lambda_1 < \lambda_2 . \tag{14}$$

This test is identical with Test 1 when the identities of the experiments are interchanged.

$$\text{Test 3.} \qquad H_0 : \lambda_1 = \lambda_2 = \lambda, \qquad H_{a_3} : \lambda_1 \neq \lambda_2 . \tag{15}$$

The two-sided test is of course similar to Tests 1 and 2 and would be used when departures from H_0 in either direction are of interest. For significance level 2α, reject H_0 in favor of H_{a_3} when $w \geq w_0(\alpha)$ or when $1/w \geq w_0(\alpha)$.

(ii) Tests of hypotheses on correlation

$$\text{Test 4.} \qquad H_0 : \rho_1^2 = \rho_2^2 , \qquad H_{a_4} : \rho_1^2 > \rho_2^2 . \tag{16}$$

$$\text{Test 5.} \qquad H_0 : \rho_1^2 = \rho_2^2 , \qquad H_{a_5} : \rho_1^2 < \rho_2^2 . \tag{17}$$

$$\text{Test 6.} \qquad H_0 : \rho_1^2 = \rho_2^2 , \qquad H_{a_6} : \rho_1^2 \neq \rho_2^2 . \tag{18}$$

Tests 4, 5, and 6 may be associated with Tests 1, 2, and 3 respectively when we redefine

$$w = R_1^2(1 - R_2^2)/R_2^2(1 - R_1^2) \tag{19}$$

with the association of the parameters in (12). With this new definition of w, the test procedures are exactly those used in (*i*) above for tests of hypotheses on sensitivity. It is of course necessary that $(a + b)$ in (12) be constant for the two regression experiments.

(iii) Analysis

The following easy basic steps are required in the tests outlined above.

1. Do analyses of variance for both experiments yielding F-statistics valid for treatment comparisons, each with $2a$ and $2b$ degrees of freedom. [When analysis of variance for multiple regression is used, $F = bR^2/a(1 - R^2)$].

2. Estimate the parameters λ_1 and λ_2 pertaining to the two experiments.

(*a*) In Model I of the analysis of variance, estimate λ for each experiment by taking

$$\hat{\lambda} = a\left[\frac{s_t^2}{s_e^2} - 1\right] \tag{20}$$

where s_t^2 and s_e^2 are respectively mean squares for treatments and for error. The mean of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ is heuristically taken to be the "best available" single estimate* of the population parameter λ.

(*b*) In multiple regression, Snedecor [1950, p. 348] gives

$$R_A^2 = 1 - (1 - R^2)(a + b)/b \tag{21}$$

*The conservative statistician might take $\hat{\lambda}$ to be the smaller of $\hat{\lambda}_1$ and $\hat{\lambda}_2$; the ultra-conservative statistician might take $\hat{\lambda} = 0$. For moderately large a and b, $w_0(\alpha)$ does not change much with changes in λ.

in our notation and R_A^2 may be taken as an estimate of ρ^2 required to estimate λ in (12). This is however equivalent to taking

$$\hat{\lambda} = a\left[\frac{s^2_{\text{reg}}}{s^2_{\text{res}}} - 1\right] \tag{22}$$

similar to (20) where s^2_{reg} and s^2_{res} are respectively mean squares for regression and residual (error) variation in the analysis of variance for regression. Then again estimates of λ_1 and λ_2 are obtained from the two experiments and their mean is taken as $\hat{\lambda}$.

3. Formulate the null hypothesis $H_0 : \lambda_1 = \lambda_2 = \lambda$, using $\hat{\lambda}$ for λ, and the suitable alternative hypothesis selected from Tests 1–6.

4. Decide on the significance level to be used, α for a one-sided test and 2α for a two-sided test.*

5. Calculate $w = F_1/F_2$.

6. Compute $a' = (a + \hat{\lambda})^2/(a + 2\hat{\lambda})$ using $\hat{\lambda}$ obtained in Step 2.

7. Obtain $w_0(\alpha)$ from Table 1, possible when $\alpha = .05$, by entering that table with a' for a and with b and by using linear interpolation in the table.

8. Compare w with $w_0(\alpha)$ rejecting H_0 in favor of H_a in accordance with the rules set forth for Tests 1–6.

NUMERICAL EXAMPLES

The following numerical examples have been chosen to illustrate applications of the method of comparing sensitivities of identical experiments in different experimental situations. In summarizing the analyses we follow the steps outlined in the preceding section.

Example 1. Discriminating Abilities of Judges.

Two judges were asked to score desirability of color of nine brands of canned tomatoes in two independent randomized block experiments. The discrete scoring scale employed had values ranging in units of 5 from 0 to 100. Although the scale was discrete, it will be assumed that

TABLE 2
AVERAGE COLOR SCORES FOR CANNED TOMATOES

Brand		2	5	9	3
Judge A (Exp. 1)		87.5	85.0	78.7	75.0
Judge B (Exp. 2)		75.0	68.7	65.0	60.0
Brand	**1**	**7**	**4**	**8**	**6**
Judge A	61.3	50.0	48.7	48.7	47.5
Judge B	52.5	45.0	40.0	40.0	38.7

*At the present time we are limited to taking $\alpha = .05$ due to the unavailability of other tables.

analysis of variance applies and the observed scores were used in such analyses. Judge A was to be used on a scoring panel if he demonstrated sensitivity to color differences as well as the experienced Judge B.

The nine brands of canned tomatoes with the average scores accorded each, averaged over four replications, for each judge are given in Table 2.

The analysis for comparison of the sensitivities (discriminating abilities) of the two judges now follows:

1. Analyses of variance are given in Table 3. $F_1 = 19.14$ and $F_2 = 5.42$ while $2a = 8$, $2b = 24$, $a = 4$, $b = 12$.

TABLE 3

ANALYSES OF VARIANCE FOR THE TWO JUDGES

Source	d.f.	Mean Square	F
	Judge A (Experiment 1)		
Blocks	3	167.6	
Treatments	8	1134.1	19.14**
Error	24	59.3	
	Judge B (Experiment 2)		
Blocks	3	346.3	
Treatments	8	761.6	5.42**
Error	24	140.6	

**Significant at the .01 level of significance.

2. We use method (a), equation (20).

$$\hat{\lambda}_1 = 4\left[\frac{1134.1}{59.3} - 1\right] = 72.5.$$

$$\hat{\lambda}_2 = 4\left[\frac{761.6}{140.6} - 1\right] = 17.7.$$

We take $\hat{\lambda} = \frac{1}{2}(17.7 + 72.5) = 45.1$.

3. The hypotheses are $H_0 : \lambda_1 = \lambda_2 = \lambda$, $H_{a_1} : \lambda_1 > \lambda_2$ as described in Test 1.

4. It is specified that $\alpha = .05$.

5. $w = 19.14/5.42 = 3.53$.

6. $a' = (4 + 45.1)^2/(4 + 90.2) = 25.6$ from (9).

7. $w_0(.05) < 2.87$, obvious from Table 1 with $a' = 25.6$, $b = 12$ although a' exceeds the tabled values of a in Table 1.

8. H_0 is rejected in favor of H_{a_1} on the basis of the test for $w > w_0(.05)$. This implies that Judge A is a good judge of color differences even in comparison with experienced Judge B and Judge A may be included on the scoring panel.

Note that if we had taken $\hat{\lambda} = \hat{\lambda}_2 = 17.7$, we would have obtained $a' = (4 + 17.7)^2/(4 + 35.4) = 12.0$ and $w_0(.05) \approx 2.60$. Our conclusions would of course be unchanged.

Examination of Table 2 suggests that the two judges are not using the scoring scale in quite the same way but that the difference is largely due to location and scale-unit differences. There is no evidence of judge by treatment interaction of the sort that would make interpretation of the w-test used here difficult. In most cases it would be necessary to train Judge A in use of the scoring scale consistent with that of the other panel members before he may be included on the panel.

Example 2. Sensitivity of Pasture Yields to Irrigation.

The yields of a mixture of Orchard Grass and Ladino Clover subjected to eight different treatments (eight fertilizers) were obtained from plots which were either irrigated or not irrigated. The experiments relating to irrigation and no irrigation were independent and each contained four replications. The average yields in pounds per acre for treatments A to H, averaged over the four replications, are given in Table 4. An objective was to learn whether the effects of differences in fertilization would be more or less pronounced under irrigation than under no irrigation.

Many irrigation experiments of the kind described have been conducted at the Virginia Agricultural Experiment Station and little treatment by irrigation interaction has been obtained that is of a fundamental nature (that is, that cannot be explained by a shift in treatment locations and a magnification of treatment differences due to irrigation). If, for example, we plot pairs of observations in columns of Table 4, we find the points so obtained very nearly collinear and this is a crude but adequate check on the explanation of the treatment by irrigation interaction.

Sometimes split-plot designs are used in this sort of irrigation experiment, but this is not usually acceptable since error variances for irrigated plots are often eight to ten times as large (only 2.7 times as large in our example) as for plots without irrigation. This same heterogeneity of error variances also largely precludes the possibility of

TABLE 4

Mean Yields of Orchard Grass and Ladino Clover in Lbs/Acre Under Fertilizer Treatments

Treatment	A	B	C	D	E	F	G	H
No irrigation (Exp. 1)	499	627	642	756	884	494	597	377
Irrigation (Exp. 2)	831	836	931	1137	1346	730	903	583

looking at tables of means such as Table 4 in order to compare the sensitivities of the two growing conditions to treatment differences. The split-plot design is not necessary (whole plot treatments were "irrigation" and "no irrigation") for in most experiments in most years huge differences between irrigation and no irrigation are found.

The analysis for comparison of the sensitivities of no-irrigation and irrigation experiments to differences in fertilizer treatments is now set out.

1. $F_1 = 8.86$, $F_2 = 7.10$, $a = 3.5$, and $b = 10.5$ from analyses of variance in Table 5.

2. $\hat{\lambda}_1 = 3.5(8.86 - 1) = 27.5$ and $\hat{\lambda}_2 = 3.5(7.10 - 1) = 21.4$ and we take $\hat{\lambda} = \frac{1}{2}(27.5 + 21.4) = 24.4$ using (20) and Step 2.

3. $H_0 : \lambda_1 = \lambda_2 = \lambda$ and $H_{a_2} : \lambda_1 \neq \lambda_2$ were chosen in accordance with the objective of the experiments.

4. We take $\alpha = .05$ and the significance level is .10.

5. $w = 8.86/7.10 = 1.25$.

6. $a' = (3.5 + 24.4)^2/(3.5 + 48.8) = 14.9$ from (9).

7. We use the result that $G(w_0 ; a, b, 0) = G(w_0 ; b, a, 0)$ and enter Table 1 with $a = 10.5$ (in our experiments $b = 10.5$) and with $b = 14.9$ (we obtained $a' = 14.9$). Then we find that $2.36 < w_0(.05) < 2.63$.

TABLE 5

Analyses of Variance for the Two Fertilizer Experiments

Source	d.f.	Mean Square	F
	Experiment 1—No Irrigation		
Replications	3	17,899	
Treatments	7	102,104	8.86**
Error	21	11,517	
	Experiment 2—Irrigation		
Replications	3	93,155	
Treatments	7	224,462	7.10**
Error	21	31,599	

**Significant at the .01 level of significance.

8. w is not significant in comparison with $w_0(.05)$ nor is $1/w$. It does not appear that these irrigation and no-irrigation experiments differ much in sensitivity to fertilizer treatment effects. (While we took, $\hat{\lambda} = 24.4$, the same conclusion would have been reached for other values of λ in a fairly wide interval about 24.4.)

Example 3. A Comparison of Two Multiple Correlation Coefficients.

In research in chemical engineering on an extraction mechanism, the rate of transfer of one component of a two-component, two-phase

system (water and a solvent) across the interface was measured.* This rate of transfer was correlated with the physical properties of the phases. For any one system, log $(k_s d/D_w)$ or y was assumed to be linearly related to log $(G_s d/\mu_s)$ or x_1 and log $(G_w d/\mu_w)$ or x_2 as a result of a multiplicative model based on theoretical knowledge of the mechanism. The symbols used are:

k_s : The film coefficient for transfer of water into the solvent phase, lb. moles per hr., sq. ft. unit concentration difference.
d : The diameter of the extraction tube, ft.
D_w : The diffusivity of water into solvent, sq. ft. per hr.
G_s : The mass velocity of the solvent phase, lb. per hr., sq. ft.
G_w : The mass velocity of the water phase, lb. per hr., sq. ft.
μ_s : The viscosity of the solvent phase, lb. per ft., hr.
μ_w : The viscosity of the water phase, lb. per ft., hr.

For this example, we consider regression equations of the form

$$y = a + b_1 x_1 + b_2 x_2$$

for two solvents, cyclohexanol and methyl ethyl ketone. The independent variables, x_1 and x_2 , are measurable with small errors that are insignificant in comparison with the residual variability of y. Accordingly, the regression equations and the resulting estimates of the multiple correlation coefficients are appropriate to illustrate the comparison described in this paper. We are interested in comparing the dependencies of y, essentially a transfer rate, on the physical properties of the phases for the two different solvents. The experimental results and the statistical analysis are now set forth.

The regression equations are:

1. Cyclohexanol,

$$y = 1.0986 + 1.0931 x_1 + 0.1736 x_2$$

and

2. Methyl ethyl ketone,

$$y = -2.8715 + 0.8712 x_1 + 0.5234 x_2 \ .$$

The corresponding squared multiple correlation coefficients are $R_1^2 = 0.571$ and $R_2^2 = 0.486$ with 2 and 34 and 2 and 31 degrees of freedom respectively.

The two regression equations look rather different but the objective

*We are indebted to Professor N. F. Murphy of the Department of Chemical Engineering, Virginia Polytechnic Institute, for permission to use his data and for advice on the presentation of this example. The regression equations in this example were calculated from work on liquid-liquid extraction given by Murphy, Lastovica, and Skrzec [1956] which contains similar, but somewhat more complex, regression analyses.

is simply to compare values of ρ^2 as measures of the effectiveness of the regressions. This can be done here since the degrees of freedom are almost the same for the two studies and this is required in view of the definition of λ in (12).

The analysis is:

1. $R_1^2/(1 - R_1^2) = 1.3310$, $R_2^2/(1 - R_2)^2 = 0.9455$ with $a = 1$ and $b_1 = 17$, $b_2 = 15.5$. We take $\bar{b} = (17 + 15.5)/2 = 16.25$, the average of the values of b for R_1^2 and R_2^2.

2. Using Step 2(b), formula (21), we have

$$\hat{\rho}_1^2 = 1 - (1 - 0.571)(1 + 17)/17 = 0.546,$$

$$\hat{\rho}_2^2 = 1 - (1 - 0.486)(1 + 15.5)/15.5 = 0.453.$$

Then, from (12),

$$\hat{\lambda}_1 = (1 + 17)(0.546)/(1 - 0.546) = 21.647$$

and

$$\hat{\lambda}_2 = (1 + 15.5)(0.453)/(1 - 0.453) = 13.665.$$

We take

$$\hat{\lambda} = \tfrac{1}{2}(21.647 + 13.665) = 17.66.$$

3. The hypotheses are $H_0 : \rho_1^2 = \rho_2^2$ and $H_{a_0} : \rho_1^2 \neq \rho_2^2$.

4. $\alpha = 0.05$; the significance level is $2\alpha = 0.10$.

5. $w = 1.3310/0.9455 = 1.408$ from (19).

6. $a' = (1 + 17.66)^2/[1 + 2(17.66)] = 9.59$.

7. It is evident from Table 1 that $w_0(.05) \approx 2.67$ and certainly $w_0(.05) > 2.17$ when we enter that table with $a = 9.59$ and $b = \bar{b} = 16.25$.

8. The observed value of w, 1.408, is exceeded by $w_0(.05) > 2.17$ and we do not reject $H_0 : \rho_1^2 = \rho_2^2$ formulated in Step 3 above.

DISCUSSION AND SUMMARY

We have illustrated a method of comparing the sensitivities of two identically designed experiments and of comparing the squares of two multiple correlation coefficients from comparable regression analyses. In each case, the problem reduced to a test of the hypothesis of equality of two parameters of noncentrality in two noncentral F-distributions. This is satisfactory when the number of replications in each treatment mean is the same in each analysis of variance and when each regression equation is based on the same number of observations.

Examples where these methods may be used have been pointed out in a general way in a number of areas of research. Three particular numerical examples are given illustrating the application of the steps

for analysis specifically set forth. Some discussion of the care needed in interpreting these analyses is given and in particular it is pointed out that treatment by experiment interactions of a fundamental nature may lead to difficulties of interpretation. The table available for the tests of significance required in analyses is somewhat limited (this is indicated from the way the examples are handled) but Schumann is planning to extend these tables and to prepare tables for values of α, the significance level in one-sided tests, of .025, 0.01, and .005. Nevertheless, the available table with $\alpha = .05$ does provide a means of making the tests that are often desired in experimental work.

There are some situations wherein the sensitivities of experiments that do not meet the requirements on similarity set forth should be compared. The practical situations seem to be where the same treatments are in both experiments but where the experiments are based on different numbers of replications or on somewhat different experimental designs. A crude method of obtaining an approximation to the critical value of w in that situation based on the use of Table 1 has been suggested by Schumann and Bradley [1957] but its accuracy is not known and it is not repeated here.

In conclusion, D. E. W. Schumann wishes to express his appreciation of a grant from the South African Council of Industrial and Scientific Research, Pretoria, South Africa, and for assistance from Fulbright and U. S. Department of State grants. These grants and the assistance of the Agricultural Research Service, U.S.D.A., made this work possible.

We are grateful to Mrs. Sally Arnold, who did much of the computing, and to Mrs. E. P. Ellison, Miss Marianne Byrd, and Mrs. J. D. Hall for clerical assistance.

REFERENCES

Baten, W. D. [1946]. Organoleptic tests pertaining to apples and pears. *Food Research, 11*, 84–94.

Cochran, W. G. [1943]. The comparison of different scales of measurement for experimental results. *Ann. Math. Stat., 14*, 205–216.

Hopkins, J. W. [1950]. A procedure for quantifying subjective appraisals of odor, flavor and texture of foodstuffs. *Biometrics, 6*, 1–16.

Hopkins, J. W. [1953]. Laboratory flavor scoring: Two experiments in incomplete blocks. *Biometrics, 9*, 1–21.

Murphy, N. F., Lastovica, J. E., and Skrzec, A. E. [1956]. Mass transfer in a horizontal liquid-liquid extraction tube. To be published in *J. Amer. Inst. Chem. Eng.*

Schumann, D. E. W. and Bradley, R. A. [1957]. The comparison of the sensitivities of similar experiments: Theory. *Ann. Math. Stat., 28.*

Schumann, D. E. W. [1956]. *The comparison of the sensitivities of experiments using different scales of measurement.* Ph.D. dissertation, Virginia Polytechnic Institute Library, Blacksburg, Virginia.

ERROR RATES AND SAMPLE SIZES FOR RANGE TESTS IN MULTIPLE COMPARISONS

H. Leon Harter
Wright Air Development Center
Wright-Patterson Air Force Base, Ohio, U. S. A.

1. *Introduction*

The problem of multiple comparisons has recently aroused a great deal of interest among statisticians. The basic F-test in an analysis of variance determines whether there is a significant difference among a group of means, but it cannot tell which means differ significantly from which others. The latter is often what the investigator really wants to know. Various multiple comparisons tests, including the range tests discussed in this paper, have been proposed. A study is made here of the error rates, α and β, and their relation to sample size, N, for three fixed range tests and three multiple range tests.

Let it be required to test the significance of the range of p out of m ordered means of samples of size N, where $p = 2, 3, \cdots, m$. For the fixed range tests, the critical range for a particular Type I error rate, α, depends only on m and N; for the multiple range tests, it depends also on p. In fact, for fixed m and N, it is a non-decreasing function of p. The range of p means is said to be significant or non-significant (at the α level) according as it does or does not exceed the critical range, except that the range of p means is automatically non-significant if these p means constitute a subgroup of a larger group whose range is non-significant. Because of this exception, a multiple range test starts with a test on all m means. If the range of all m means is found to be significant, then tests are performed on the ranges of $(m-1)$ successive means, $(m-2)$ successive means, and so on until significant differences are no longer found.

2. *Significance vs. Confidence*

Though the emphasis here is upon tests of significance of size α for the difference between sample means, each of the above methods can also be used for setting $100\ (1-\alpha)\%$ confidence limits on the

difference between population means. The choice between significance tests and confidence limits depends upon the experimental situation and upon the experimenter's philosophy. Many statisticians, including Tukey [1953], prefer confidence procedures in practically all cases, and the author is convinced that there is much merit in this position. The results of this paper, however, will be stated in terms of significance. The reader who prefers confidence procedures can easily translate from the language of significance to that of confidence.

3. *The Fixed Range Tests*

The three fixed range tests which will be considered are the least significant difference (LSD) test based upon the work of "Student" [1908], a modification of this proposed by Fisher [1935], and the studentized range test proposed by Tukey [1952]. They will be designated by code letters L, E and W respectively.

The differences LSD, WSD and ESD between any two out of m means of samples of size N required for significance (at the α level) by the LSD test, Tukey's studentized range test, and Fisher's test, respectively, are given by

$$\text{LSD} = t(\alpha, \nu)s_{\bar{d}} = q(\alpha, 2, \nu)s_{\bar{x}} , \tag{1}$$

$$\text{WSD} = q(\alpha, m, \nu)s_{\bar{x}} , \tag{2}$$

$$\text{ESD} = t(\alpha/C_2^m , \nu)s_{\bar{d}} = q(\alpha/C_2^m , 2, \nu)s_{\bar{x}} , \tag{3}$$

where ν is the number of degrees of freedom for the error mean square, s^2 ; $t(\alpha, \nu)$ is the 2-tailed α point (upper $\alpha/2$ point) of Student's t with ν degrees of freedom; $q(\alpha, r, \nu)$ is the upper α point of the studentized range of r observations with ν degrees of freedom for s; $s_{\bar{x}}$ is the standard error of the mean, $s_{\bar{x}} = \sqrt{s^2/N}$; and $s_{\bar{d}}$ is the standard error of the difference between means, $s_{\bar{d}} = \sqrt{2s^2/N} = \sqrt{2}\, s_{\bar{x}}$.

Each of these three test procedures has Type I error rate α for an appropriate definition of the error rate. As a result of a test on the means of m samples each of size N, a statement of significance or nonsignificance will be made about each of the C_2^m differences between pairs of means. A wrong statement (Type I error) is a statement that two means are significantly different, when actually they are means of random samples from the same population. The appropriate error rates α_L , α_W , and α_E for the respective test types are then defined as follows: α_L is the expected proportion of wrong statements (error rate per comparison); α_W is the expected proportion of experiments with one or more wrong statements (familywise error rate); and α_E

is the expected number of wrong statements per experiment (error rate per family).

In practice, many statisticians (including the author) use the protected LSD test, in which the individual comparisons are made only if an analysis of variance yields a significant F-ratio for the comparison of all m means. The Type I error rate, α_L, which will be considered here, however, is that for the unprotected LSD test, otherwise known as the multiple t-test.

4. *The Multiple Range Tests*

The three multiple range tests which will be considered are the Newman-Keuls test, Tukey's X procedure, and Duncan's new multiple range test. They will be designated by the code letters I, X and S, respectively.

The Newman-Keuls test was first proposed by Newman [1939] and modified slightly by Keuls [1952]. The critical range of p out of m ordered means of samples each of size N is based on the studentized range of p observations instead of on the studentized range of m observations as in Tukey's studentized range test. It can easily be seen that the critical range ISD for the Newman-Keuls test is intermediate between the LSD and the WSD, except that it is the same as the LSD for $p = 2$ and the same as the WSD for $p = m$.

Tukey's X procedure was proposed by Tukey [1953] as a compromise between the Newman-Keuls test and Tukey's studentized range test. The critical range is the arithmetic mean of the critical ranges for those two tests. For fixed m and N, the smallest critical range (a value midway between the LSD and the WSD) occurs for $p = 2$, and the largest critical range (equal to the WSD) occurs for $p = m$.

Duncan's new multiple range test, proposed by Duncan [1953, 1955], makes use of special protection levels based upon degrees of freedom. Let $\gamma_{2,\alpha} = 1 - \alpha$ be the protection level for testing the significance of difference between two means, that is, the probability that a significant difference between sample means will not be found if the population means are equal. Duncan reasons that one has $(p-1)$ degrees of freedom for testing p means, and hence one may make $(p-1)$ independent tests, each with protection level $\gamma_{2,\alpha}$. Hence the joint protection level is

$$\gamma_{p,\alpha} = \gamma_{2,\alpha}^{p-1} = (1 - \alpha)^{p-1}; \tag{4}$$

that is, the probability that one finds no significant differences in making $(p-1)$ independent tests, each at protection level $\gamma_{2,\alpha}$, is

$\gamma_{2,\alpha}^{p-1}$. Thus Duncan's new multiple range test is based upon protection levels $\gamma_{p,\alpha}$ for tests on p means. The critical range SSD is a minimum (equal to the LSD) for $p = 2$, and a maximum for $p = m$.

The critical ranges ISD, XSD, and SSD for p out of m ordered means of samples of size N required for significance (at the α level) by the Newman-Keuls test, Tukey's X procedure, and Duncan's new multiple range test, respectively, are given by

$$\text{ISD} = q(\alpha, p, \nu)s_{\bar{x}} , \tag{5}$$

$$\text{XSD} = (1/2)[q(\alpha, p, \nu) + q(\alpha, m, \nu)]s_{\bar{x}} , \tag{6}$$

$$\text{SSD} = R(p, \nu, \gamma_{p,\alpha})s_{\bar{x}} , \tag{7}$$

where $R(p, \nu, \gamma_{p,\alpha})$ is the studentized range of p observations with ν degrees of freedom for s and protection level $\gamma_{p,\alpha}$ given by equation (4), and where the other symbols are defined as in the previous section.

5. *Example*

At this point an example will serve to illustrate the various test procedures. For the sake of simplicity, a single-classification experiment involving comparisons of five means will be considered. The data represent the weight loss (grams) in four days of rats on five diets (no food and 8 calories per rat per day of four foods). In order to reduce experimental error, only male rats were used, all rats were of the same strain, and their initial weights were nearly constant. All one hundred rats received water ad libitum, but they were randomly divided into five groups of twenty rats each, one group on each of the five diets.

An analysis of variance of the data yields the following results:

Source of Variation	D.f.	S.s.	M.s.	F
Among diets	4	3012.70	753.18	38.14**
Between diets (error)	95	1876.30	19.75	
Total	99	4889.00		

Thus there is a highly significant difference among mean weight losses of the rats on the five diets. This information is quite interesting to the experimenter, but what he really wishes to know is which diets are significantly better than which others. In order to obtain this information, it is necessary to apply a multiple comparisons test to the means:

Purina Laboratory Chow	Casein	Dextrose	Olive Oil	No Food
23.70	24.90	28.25	29.40	39.25

The standard error of the mean, $s_{\bar{x}}$, and the standard error of the difference between means, $s_{\bar{d}}$, are found to be 0.994 and 1.405, respectively, based upon the internal estimate of variance, $s^2 = 19.75$, with 95 degrees of freedom. The factors by which these must be multiplied to obtain the critical ranges for significance at the 5% and 1% levels can be found in tables. Tables of the critical values of the Student t distribution can be found in almost any statistics textbook. The 5% and 1% points of the studentized range have been tabulated by Pearson and Hartley [1954], and the 5% and 1% points of $R(p, \nu, \gamma_{p,\alpha})$ by Beyer [1953] and Duncan [1955]. All of these tables have been reproduced by Federer [1955], who gives the only textbook treatment of the subject of multiple comparisons of which the author is aware. The factors by which $s_{\bar{x}} = 0.994$ must be multiplied to obtain the critical ranges for the various tests are as follows:

Code Letter	$\alpha = 0.05$ $\nu = 95$				$\alpha = 0.01$			
	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 2$	$p = 3$	$p = 4$	$p = 5$
L	2.81	2.81	2.81	2.81	3.72	3.72	3.72	3.72
S	2.81	2.96	3.06	3.12	3.72	3.87	3.98	4.07
I	2.81	3.37	3.70	3.94	3.72	4.22	4.53	4.74
X	3.38	3.66	3.82	3.94	4.23	4.48	4.64	4.74
W	3.94	3.94	3.94	3.94	4.74	4.74	4.74	4.74
E	4.05	4.05	4.05	4.05	4.80	4.80	4.80	4.80

Multiplying the standard error of the mean $s_{\bar{x}} = 0.994$ by these factors, one finds that the critical ranges for the various tests are as follows:

Code Letter	$\alpha = 0.05$				$\alpha = 0.01$			
	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 2$	$p = 3$	$p = 4$	$p = 5$
L	2.79	2.79	2.79	2.79	3.70	3.70	3.70	3.70
S	2.79	2.94	3.04	3.10	3.70	3.85	3.96	4.05
I	2.79	3.35	3.68	3.92	3.70	4.19	4.50	4.71
X	3.36	3.64	3.80	3.92	4.20	4.45	4.61	4.71
W	3.92	3.92	3.92	3.92	4.71	4.71	4.71	4.71
E	4.03	4.03	4.03	4.03	4.77	4.77	4.77	4.77

Then the results of the various tests can be summarized as follows, where two means are or are not significantly different at the level α according as they are not or are underscored by the same line:

Test	$\alpha = 0.05$					$\alpha = 0.01$				
LSD	23.70	24.90	28.25	29.40	39.25	23.70	24.90	28.25	29.40	39.25
SSD	23.70	24.90	28.25	29.40	39.25	23.70	24.90	28.25	29.40	39.25
ISD	23.70	24.90	28.25	29.40	39.25	23.70	24.90	28.25	29.40	39.25
XSD	23.70	24.90	28.25	29.40	39.25	23.70	24.90	28.25	29.40	39.25
WSD	23.70	24.90	28.25	29.40	39.25	23.70	24.90	28.25	29.40	39.25
ESD	23.70	24.90	28.25	29.40	39.25	23.70	24.90	28.25	29.40	39.25

Thus, at the 5% level, the only disagreement is as to whether the means for casein and dextrose differ significantly. At the 1% level, there are disagreements in regard to the following differences: (1) casein and olive oil and (2) Purina laboratory chow and dextrose. For experiments involving a larger number of means, one would expect a greater number of disagreements.

6. *Type I Error Rates*

The Type I error rate, α, is defined differently for the various tests. The Type I error rate (as defined for a particular test) is denoted by α with the code letter for that test as a subscript; for example, α_L is the Type I error rate as defined by the LSD test. A comparison of the error rates is in order.

By comparison of (1) with (2), one sees that a WSD with Type I error rate α_W is equivalent to an LSD with Type I error rate α_L, where

$$t(\alpha_L, \nu) \equiv q(\alpha_L, 2, \nu)/\sqrt{2} = q(\alpha_W, m, \nu)/\sqrt{2}. \tag{8}$$

Similarly, by comparison of (1) with (3), one sees that an ESD with Type I error rate α_E is equivalent to an LSD with Type I error rate α_L, where

$$\alpha_L = \alpha_E/C_2^m. \tag{9}$$

Let $\alpha_L \mid \alpha_M$ be the value of α_L corresponding to a given α_M, where M may be any one of the letters W, E, I, X, or S. Table 1A gives $\alpha_L \mid \alpha_W$ corresponding to $\alpha_W = 0.05$, 0.01 for $m = 2(1)10(2)20$, and $N = 2, 3, 4, 6, 10, 16, 25, \infty$; also $\alpha_L \mid \alpha_E$ corresponding to $\alpha_E = 0.05$,

0.01 for the same values of m. (The relation between α_L and α_E is independent of N.) This table, as well as others which will follow, has been computed for a single-classification design, and the results are slightly different for randomized blocks or other experimental designs, due to a slight reduction in the number of degrees of freedom.

In testing the significance of the range of p out of m ordered means of samples of size N by one of the multiple range tests with Type I error rate α (as defined for that particular test), the critical range is a function of m, N and p. If α, m and N are fixed, the critical range is a function of p only. Each critical range, for fixed m and N, corresponds to a particular value of the Type I error rate α_L ; the larger the critical range, the smaller the value of α_L to which it corresponds. Since the critical range is a non-decreasing function of p, the maximum value of α_L corresponding to a fixed α for the multiple range test under consideration, with m and N also fixed, occurs for $p = 2$, and the minimum for $p = m$.

For $p = 2$, the Newman-Keuls test and Duncan's new multiple range test have critical ranges equal to the LSD; hence $(\alpha_L)_{\max} \mid \alpha_I$ is equal to α_I and $(\alpha_L)_{\max} \mid \alpha_S$ is equal to α_S .

For $p = m$, the Newman-Keuls test and Tukey's X procedure have critical ranges equal to the WSD; hence $(\alpha_L)_{\min} \mid \alpha_I$ is equal to $\alpha_L \mid \alpha_W = \alpha_I$, and $(\alpha_L)_{\min} \mid \alpha_X$ is equal to $\alpha_L \mid \alpha_W = \alpha_X$. The values of $\alpha_L \mid \alpha_W = 0.05, 0.01$ have been given in Table 1A.

Table 1B gives $(\alpha_L)_{\max} \mid \alpha_X$ corresponding to $\alpha_X = 0.05, 0.01$ and Table 1C gives $(\alpha_L)_{\min} \mid \alpha_S$ corresponding to $\alpha_S = 0.05, 0.01$, for the same values of m and N as in Table 1A.

7. *Type II and Type III Error Rates*

Let m_U and m_L be the two extreme values of p out of m ordered means of samples of size N. Let μ_U and μ_L be the population means corresponding to m_U and m_L respectively. Let the subscripts U and L be so assigned that $\mu_U \geq \mu_L$. (This does not necessarily imply that $m_U \geq m_L$.) It should be noted that, while exactly $(p - 2)$ of the $(m - 2)$ other sample means lie between m_U and m_L , the number of other population means lying between μ_U and μ_L may be less than, equal to, or greater than $(p - 2)$. Now define δ and δ' by the relations

$$\delta = (\mu_U - \mu_L)/\sigma, \tag{10}$$

$$\delta' = (\mu_U - \mu_L)/s, \tag{10'}$$

where σ is the common population standard deviation and s is the pooled estimate obtained from the data.

If the conclusion is drawn that m_U and m_L are not significantly

different when actually $\mu_U > \mu_L$, this constitutes an error of Type II. If m_L is found to be significantly greater than m_U when actually $\mu_U > \mu_L$, this will be called an error of Type III. Figure 1 shows how such an

TABLE 1

COMPARISON OF TYPE I ERROR RATES FOR SINGLE CLASSIFICATION

A. Values of α_L Corresponding to $\alpha_W = 0.05, 0.01$ and to $\alpha_E = 0.05, 0.01$

m	$\alpha_L \mid \alpha_W = 0.05$								$\alpha_L \mid \alpha_E = 0.05$
	$N=2$	$N=3$	$N=4$	$N=6$	$N=10$	$N=16$	$N=25$	$N\to\infty$	(all N's)
2	.05000	.05000	.05000	.05000	.05000	.05000	.05000	.05000	.05000
3	.02506	.02197	.02093	.02028	.01964	.01943	.01923	.01911	.01667
4	.01517	.01254	.01170	.01104	.01070	.01040	.01032	.01023	.00833
5	.01022	.00818	.00746	.00699	.00670	.00654	.00646	.00634	.00500
6	.00727	.00569	.00524	.00486	.00465	.00450	.00444	.00438	.00333
7	.00542	.00418	.00382	.00356	.00341	.00331	.00327	.00319	.00238
8	.00418	.00319	.00294	.00272	.00260	.00255	.00249	.00242	.00179
9	.00330	.00251	.00231	.00216	.00202	.00195	.00194	.00191	.00139
10	.00269	.00204	.00185	.00173	.00165	.00162	.00161	.00157	.00111
12	.00184	.00141	.00129	.00119	.00114	.00112	.00111	.00109	.00076
14	.00135	.00103	.00094	.00089	.00085	.00084	.00082	.00080	.00055
16	.00103	.00079	.00072	.00067	.00065	.00064	.00063	.00062	.00042
18	.00080	.00062	.00057	.00053	.00051	.00050	.00050	.00049	.00033
20	.00064	.00050	.00047	.00043	.00041	.00041	.00040	.00040	.00026
m	$\alpha_L \mid \alpha_W = 0.01$								$\alpha_L \mid \alpha_E = 0.01$
	$N=2$	$N=3$	$N=4$	$N=6$	$N=10$	$N=16$	$N=25$	$N\to\infty$	(all N's)
2	.01000	.01000	.01000	.01000	.01000	.01000	.01000	.01000	.01000
3	.00497	.00420	.00397	.00383	.00372	.00364	.00360	.00358	.00333
4	.00291	.00233	.00215	.00201	.00192	.00188	.00187	.00186	.00167
5	.00191	.00146	.00133	.00125	.00119	.00117	.00116	.00114	.00100
6	.00134	.00101	.00092	.00085	.00081	.00080	.00079	.00076	.00067
7	.00098	.00073	.00066	.00062	.00058	.00057	.00057	.00056	.00048
8	.00074	.00055	.00050	.00047	.00044	.00043	.00043	.00042	.00036
9	.00058	.00043	.00039	.00036	.00035	.00034	.00034	.00033	.00028
10	.00047	.00034	.00031	.00029	.00028	.00027	.00027	.00026	.00022
12	.00031	.00023	.00021	.00020	.00019	.00019	.00018	.00018	.00015
14	.00022	.00017	.00016	.00015	.00014	.00014	.00014	.00013	.00011
16	.00017	.00013	.00012	.00011	.00011	.00011	.00011	.00010	.00008
18	.00013	.00010	.00009	.00009	.00008	.00008	.00008	.00008	.00007
20	.00010	.00008	.00007	.00007	.00007	.00007	.00007	.00006	.00005

error may arise. The sum of the Type II error rate, β, and the Type III error rate, γ, will be denoted by β', that is,

$$\beta' = \beta + \gamma. \tag{11}$$

The Type III error rate, γ, may be sizeable when δ' and N are both small. Otherwise, $\gamma \doteq 0$ and $\beta' \doteq \beta$.

TABLE 1 (continued)

B. Maximum Values of α_L Corresponding to $\alpha_X = 0.05, 0.01$

m	$(\alpha_L)_{\max} \mid \alpha_X = 0.05$							
	$N = 2$	$N = 3$	$N = 4$	$N = 6$	$N = 10$	$N = 16$	$N = 25$	$N \to \infty$
2	.0500	.0500	.0500	.0500	.0500	.0500	.0500	.0500
3	.0348	.0329	.0323	.0321	.0316	.0316	.0314	.0314
4	.0266	.0249	.0244	.0240	.0238	.0236	.0235	.0235
5	.0217	.0202	.0197	.0194	.0192	.0191	.0191	.0191
6	.0183	.0170	.0167	.0164	.0163	.0162	.0162	.0162
7	.0158	.0148	.0145	.0143	.0142	.0142	.0142	.0142
8	.0140	.0130	.0129	.0127	.0127	.0126	.0126	.0126
9	.0125	.0117	.0116	.0115	.0114	.0113	.0113	.0114
10	.0114	.0107	.0105	.0104	.0104	.0104	.0104	.0104
12	.0096	.0091	.0090	.0090	.0089	.0089	.0089	.0090
14	.0084	.0080	.0079	.0079	.0079	.0079	.0079	.0079
16	.0074	.0071	.0071	.0071	.0071	.0071	.0071	.0071
18	.0067	.0065	.0064	.0064	.0064	.0064	.0065	.0065
20	.0061	.0059	.0059	.0059	.0059	.0059	.0059	.0059
m	$(\alpha_L)_{\max} \mid \alpha_X = 0.01$							
	$N = 2$	$N = 3$	$N = 4$	$N = 6$	$N = 10$	$N = 16$	$N = 25$	$N \to \infty$
2	.0100	.0100	.0100	.0100	.0100	.0100	.0100	.0100
3	.0069	.0064	.0062	.0062	.0061	.0061	.0061	.0061
4	.0052	.0048	.0046	.0045	.0045	.0045	.0045	.0045
5	.0041	.0037	.0036	.0036	.0036	.0036	.0036	.0036
6	.0034	.0031	.0031	.0030	.0030	.0030	.0030	.0030
7	.0029	.0027	.0026	.0026	.0026	.0026	.0026	.0026
8	.0025	.0023	.0023	.0023	.0023	.0023	.0023	.0023
9	.0023	.0021	.0021	.0021	.0021	.0021	.0021	.0021
10	.0020	.0019	.0019	.0019	.0019	.0019	.0019	.0019
12	.0017	.0016	.0016	.0016	.0016	.0016	.0016	.0016
14	.0015	.0014	.0014	.0014	.0014	.0014	.0014	.0014
16	.0013	.0012	.0012	.0012	.0012	.0012	.0012	.0012
18	.0011	.0011	.0011	.0011	.0011	.0011	.0011	.0011
20	.0010	.0010	.0010	.0010	.0010	.0010	.0010	.0010

TABLE 1 (continued)

C. Minimum Values of α_L Corresponding to $\alpha_S = 0.05, 0.01$

m	$(\alpha_L)_{\min} \mid \alpha_S = 0.05$							
	$N = 2$	$N = 3$	$N = 4$	$N = 6$	$N = 10$	$N = 16$	$N = 25$	$N \to \infty$
2	.0500	.0500	.0500	.0500	.0500	.0500	.0500	.0500
3	.0500	.0446	.0425	.0411	.0400	.0394	.0392	.0389
4	.0467	.0397	.0364	.0359	.0344	.0333	.0331	.0327
5	.0424	.0357	.0334	.0319	.0305	.0299	.0295	.0289
6	.0406	.0333	.0305	.0287	.0275	.0268	.0265	.0259
7	.0379	.0310	.0283	.0266	.0255	.0249	.0246	.0241
8	.0360	.0291	.0266	.0247	.0238	.0231	.0228	.0224
9	.0345	.0276	.0251	.0232	.0222	.0218	.0216	.0212
10	.0340	.0260	.0237	.0222	.0212	.0207	.0205	.0200
12	.0308	.0239	.0217	.0203	.0193	.0189	.0186	.0182
14	.0282	.0221	.0201	.0189	.0180	.0175	.0173	.0168
16	.0263	.0208	.0189	.0177	.0169	.0165	.0163	.0159
18	.0245	.0194	.0179	.0166	.0160	.0156	.0154	.0150
20	.0234	.0186	.0170	.0159	.0151	.0147	.0145	.0141
m	$(\alpha_L)_{\min} \mid \alpha_S = 0.01$							
	$N = 2$	$N = 3$	$N = 4$	$N = 6$	$N = 10$	$N = 16$	$N = 25$	$N \to \infty$
2	.0100	.0100	.0100	.0100	.0100	.0100	.0100	.0100
3	.0092	.0080	.0075	.0075	.0074	.0074	.0074	.0072
4	.0082	.0066	.0062	.0061	.0061	.0060	.0059	.0058
5	.0072	.0057	.0055	.0053	.0052	.0050	.0050	.0049
6	.0063	.0050	.0049	.0047	.0046	.0045	.0044	.0043
7	.0054	.0045	.0044	.0042	.0040	.0040	.0039	.0038
8	.0048	.0042	.0041	.0038	.0036	.0036	.0035	.0034
9	.0043	.0039	.0038	.0036	.0034	.0033	.0033	.0032
10	.0039	.0037	.0035	.0033	.0031	.0030	.0030	.0030
12	.0035	.0033	.0031	.0029	.0027	.0027	.0026	.0026
14	.0033	.0029	.0028	.0026	.0025	.0024	.0024	.0023
16	.0031	.0027	.0025	.0024	.0023	.0022	.0022	.0021
18	.0030	.0025	.0024	.0022	.0021	.0021	.0020	.0019
20	.0029	.0024	.0022	.0020	.0019	.0019	.0019	.0018

The code letters for the various tests are used as subscripts on β and β'; for example, β_L denotes the Type II error rate for the LSD test.

The Type II and Type III error rates, β and γ, and their sum, β', depend upon α, m, N and δ'. In the case of multiple range tests, they also depend upon p; for a fixed combination of α, m, N and δ', β' is a minimum for $p = 2$ and a maximum for $p = m$. For $p = 2$, the Newman-Keuls test and Duncan's new multiple range test have critical

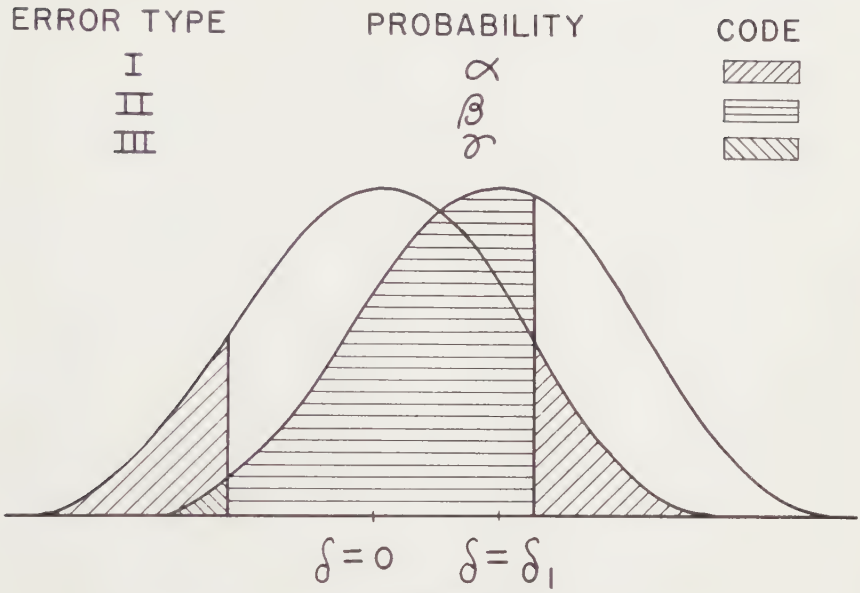


FIGURE 1
GRAPHICAL REPRESENTATION OF THREE TYPES OF ERROR

ranges equal to the LSD; hence $(\beta_I')_{\min} = (\beta_S')_{\min} = \beta_L'$. For $p = m$, the Newman-Keuls test and Tukey's X procedure have critical ranges equal to the WSD; hence $(\beta_I')_{\max} = (\beta_X')_{\max} = \beta_W'$. Table 2 gives values of β_L', $(\beta_S')_{\max}$, $(\beta_X')_{\min}$, and β_W'. For convenience, these are denoted by β_1', β_2', β_3', and β_4', respectively. Thus one has the following equations and inequalities:

$$\beta_L' = \beta_1', \tag{12}$$

$$\beta_1' \leq \beta_S' \leq \beta_2', \tag{13}$$

$$\beta_1' \leq \beta_I' \leq \beta_4', \tag{14}$$

$$\beta_3' \leq \beta_X' \leq \beta_4', \tag{15}$$

$$\beta_W' = \beta_4'. \tag{16}$$

In (13), (14), and (15), the left-hand equality holds when $p = 2$ and the right-hand equality when $p = m$. Error rates, β_E', for Fisher's test have not been computed for $m > 2$, but they are even larger than the corresponding β_W'. Hence one has the inequality

$$\beta_E' \geq \beta_4', \tag{17}$$

where the equality holds only for $m = 2$, in which case all six tests are identical.

TABLE 2

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

A. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.05$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
			$\delta' = 1.5$				$\delta' = 2.0$				$\delta' = 2.5$				$\delta' = 3.0$		
5	2	.475	.475	.475	.475	.208	.208	.208	.208	.0691	.0691	.0691	.0691	.0204	.0204	.0204	.0204
	3	.425	.466	.520	.613	.172	.199	.237	.314	.0507	.0605	.0758	.111	.0124	.0150	.0193	.0300
	4	.402	.466	.547	.685	.156	.196	.256	.385	.0427	.0573	0.818	.146	.0092	.0128	.0194	.0393
	5	.389	.471	.565	.729	.147	.199	.270	.433	.0372	.0568	.0864	.174	.0076	.0119	.0197	.0475
	6	.380	.479	.581	.760	.142	.204	.282	.472	.0348	.0577	.0909	.198	.0066	.0116	.0202	.0557
	8	.370	.489	.604	.804	.134	.210	.302	.530	.0322	.0588	.0990	.240	.0054	.0112	.0216	.0712
	10	.364	.499	.623	.833	.130	.216	.319	.575	.0303	.0606	.106	.286	.0048	.0112	.0232	.0860
	12	.359	.507	.637	.854	.128	.222	.331	.607	.0290	.0624	.113	.316	.0044	.0114	.0245	.0983
	16	.355	.521	.660	.882	.124	.232	.353	.657	.0275	.0659	.123	.351	.0039	.0118	.0272	.122
	20	.352	.533	.676	.900	.122	.240	.370	.692	.0266	.0687	.132	.387	.0036	.0127	.0295	.142
7	2	.271	.271	.271	.271	.0720	.0720	.0720	.0720	.0140	.0140	.0140	.0140	.0025	.0025	.0025	.0025
	3	.245	.278	.319	.402	.0591	.0710	.0869	.125	.0095	.0118	.0151	.0239	.0012	.0016	.0020	.0034
	4	.233	.284	.348	.481	.0532	.0715	.0978	.167	.0076	.0110	.0164	.0335	.00081	.0012	.0019	.0044
	5	.225	.295	.370	.537	.0497	.0747	.107	.203	.0066	.0110	.0176	.0428	.00061	.0011	.0019	.0055
	6	.221	.302	.387	.579	.0475	.0768	.116	.234	.0060	.0110	.0188	.0518	.00050	.0010	.0019	.0066
	8	.215	.317	.415	.641	.0448	.0817	.127	.284	.0052	.0114	.0212	.0689	.00037	.0010	.0020	.0091
	10	.212	.326	.436	.684	.0433	.0852	.138	.326	.0048	.0118	.0236	.0850	.00031	.00093	.0021	.0112
	12	.209	.335	.453	.716	.0423	.0886	.147	.359	.0045	.0122	.0252	.0994	.00027	.00093	.0023	.0143
	16	.207	.350	.479	.762	.0411	.0945	.163	.412	.0042	.0131	.0286	.125	.00023	.0010	.0026	.0196
	20	.205	.362	.499	.793	.0403	.100	.175	.474	.0040	.0140	.0317	.148	.00020	.0010	.0029	.0247
			$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$				$\delta' = 2.5$		
9	2	.500	.500	.500	.500	.152	.152	.152	.152	.0248	.0248	.0248	.0248	.0029	.0029	.0029	.0029
	3	.478	.520	.563	.645	.137	.161	.188	.250	.0197	.0246	.0307	.0467	.0017	.0023	.0029	.0049
	4	.467	.532	.598	.719	.130	.167	.212	.319	.0173	.0248	.0353	.0672	.0013	.0020	.0031	.0071
	5	.460	.547	.624	.767	.126	.176	.231	.374	.0160	.0261	.0395	.0867	.0011	.0020	.0033	.0095
	6	.456	.559	.643	.799	.123	.183	.246	.415	.0151	.0271	.0430	.114	.00094	.0020	.0035	.0118
	8	.451	5.77	.671	.842	.120	.195	.270	.480	.0142	.0293	.0492	.135	.00080	.0020	.0040	.0169
	10	.448	.590	.691	.870	.118	.204	.289	.529	.0135	.0311	.0546	.158	.00070	.0021	.0045	.0210
	12	.446	.600	.708	.890	.117	.211	.305	.568	.0131	.0326	.0594	.188	.00065	.0021	.0049	.0270
	16	.444	.616	.731	.914	.116	.223	.329	.623	.0126	.0351	.0675	.228	.00058	.0023	.0057	.0365
	20	.442	.630	.748	.930	.115	.234	.348	.664	.0123	.0378	.0742	.263	.00055	.0024	.0065	.0454

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

A. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.05$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 0.5$				$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$			
11	2	.814	.814	.814	.814	.399	.399	.399	.399	.0839	.0839	.0839	.0839	.0085	.0085	.0085	.0085
	3	.804	.832	.856	.897	.382	.423	.465	.549	.0755	.0907	.108	.151	.0064	.0082	.0105	.0082
	4	.799	.843	.877	.930	.374	.439	.502	.630	.0711	.0962	.125	.204	.0055	.0083	.0122	.0083
	5	.796	.852	.891	.948	.369	.454	.530	.686	.0688	.102	.139	.248	.0050	.0087	.0138	.0087
	6	.794	.860	.901	.959	.366	.467	.550	.724	.0672	.107	.150	.283	.0046	.0091	.0151	.0091
	8	.792	.871	.914	.972	.362	.486	.581	.777	.0653	.115	.168	.343	.0042	.0099	.0177	.0099
	10	.790	.879	.923	.979	.360	.501	.604	.813	.0644	.122	.183	.390	.0040	.0106	.0200	.0106
	12	.789	.884	.930	.984	.358	.511	.623	.839	.0634	.127	.196	.429	.0039	.0111	.0222	.0111
	16	.788	.892	.939	.989	.356	.528	.649	.871	.0624	.136	.215	.485	.0037	.0121	.0257	.0121
	20	.787	.899	.946	.992	.355	.543	.669	.893	.0619	.144	.231	.530	.0036	.0132	.0289	.0132
13	2	.781	.781	.781	.781	.316	.316	.316	.316	.0455	.0455	.0455	.0455	.0029	.0029	.0029	.0029
	3	.772	.802	.829	.875	.302	.340	.378	.458	.0404	.0500	.0606	.0879	.0020	.0027	.0035	.0059
	4	.767	.816	.853	.914	.296	.358	.416	.545	.0380	.0539	.0718	.125	.0017	.0027	.0041	.0094
	5	.764	.826	.869	.936	.292	.371	.444	.604	.0365	.0570	.0809	.158	.0015	.0028	.0046	.0130
	6	.762	.835	.880	.949	.290	.385	.465	.647	.0356	.0607	.0885	.186	.0014	.0029	.0051	.0167
	8	.760	.847	.896	.964	.287	.404	.497	.708	.0345	.0662	.101	.235	.0013	.0032	.0060	.0242
	10	.759	.856	.906	.973	.285	.417	.521	.754	.0338	.0703	.112	.275	.0012	.0034	.0069	.0317
	12	.758	.862	.915	.979	.284	.429	.540	.780	.0333	.0741	.121	.309	.0011	.0036	.0078	.0391
	16	.757	.872	.925	.985	.282	.447	.568	.821	.0328	.0803	.136	.362	.0010	.0040	.0092	.0527
	20	.756	.880	.933	.989	.281	.462	.590	.849	.0324	.0859	.148	.406	.0010	.0044	.0106	.0658
16	2	.733	.733	.733	.733	.219	.219	.219	.219	.0178	.0178	.0178	.0178	.00055	.00055	.00055	.00055
	3	.724	.758	.787	.841	.210	.242	.273	.344	.0155	.0200	.0245	.0378	.00035	.00045	.00063	.0012
	4	.720	.776	.816	.888	.206	.259	.308	.428	.0144	.0216	.0298	.0578	.00027	.00047	.00074	.0019
	5	.717	.787	.835	.915	.203	.270	.333	.488	.0137	.0230	.0343	.0763	.00023	.00047	.00084	.0028
	6	.716	.797	.849	.931	.201	.283	.353	.534	.0134	.0247	.0385	.0937	.00021	.00049	.00093	.0037
	8	.714	.812	.867	.951	.199	.300	.383	.600	.0128	.0273	.0448	.124	.00017	.00053	.0011	.0057
	10	.713	.821	.880	.963	.198	.312	.407	.649	.0125	.0294	.0506	.152	.00016	.00057	.0013	.0079
	12	.712	.829	.890	.970	.197	.323	.427	.686	.0123	.0313	.0557	.177	.00015	.00061	.0015	.0101
	16	.711	.841	.903	.979	.196	.340	.455	.736	.0120	.0344	.0640	.218	.00014	.00068	.0018	.0146
	20	.710	.850	.912	.985	.195	.354	.477	.778	.0119	.0373	.0712	.259	.00014	.00076	.0021	.0200

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

A. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.05$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 0.5$				$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$			
21	2	.655	.655	.655	.655	.115	.115	.115	.115	.0035	.0035	.0035	.0035	.00003	.00003	.00003	.00003
	3	.647	.686	.719	.782	.110	.131	.152	.203	.0029	.0039	.0050	.0085	.00002	.00002	.00003	.00007
	4	.644	.707	.753	.841	.107	.143	.177	.271	.0026	.0043	.0063	.0141	.00001	.00002	.00004	.00012
	5	.642	.720	.776	.876	.106	.152	.196	.323	.0025	.0046	.0074	.0200	.00001	.00002	.00004	.00018
	6	.640	.733	.793	.898	.105	.161	.212	.367	.0023	.0050	.0085	.0261	.00001	.00002	.00004	.00025
	8	.638	.750	.816	.925	.104	.173	.237	.433	.0022	.0056	.0103	.0376	.00001	.00002	.00006	.00041
	10	.637	.763	.832	.942	.103	.184	.256	.484	.0021	.0062	.0120	.0491	.00001	.00002	.00007	.00061
	12	.637	.772	.845	.953	.103	.192	.273	.525	.0021	.0067	.0135	.0602	*	.00003	.00008	.00084
	16	.636	.786	.863	.964	.102	.205	.298	.573	.0020	.0075	.0161	.0766		.00003	.00010	.0012
	20	.635	.797	.874	.974	.102	.216	.318	.628	.0020	.0083	.0184	.0983		.00003	.00012	.0019
25	2	.596	.596	.596	.596	.0670	.0670	.0670	.0670	.00094	.00094	.00094	.00094				
	3	.589	.630	.664	.734	.0636	.0777	.0920	.129	.00074	.0010	.0013	.0024				.00001
	4	.586	.653	.703	.801	.0621	.0862	.110	.181	.00065	.0011	.0017	.0043				.00001
	5	.584	.668	.728	.841	.0612	.0925	.124	.223	.00059	.0012	.0020	.0063				.00002
	6	.583	.682	.747	.868	.0606	.0990	.136	.260	.00056	.0013	.0024	.0086				.00002
	8	.581	.701	.773	.902	.0599	.108	.155	.319	.00052	.0015	.0029	.0132				.00004
	10	.580	.714	.791	.922	.0594	.115	.170	.366	.00050	.0016	.0035	.0169			.00001	.00007
	12	.579	.725	.806	.936	.0590	.122	.184	.406	.00048	.0018	.0040	.0229			.00001	.00010
	16	.578	.741	.826	.953	.0587	.131	.204	.464	.00047	.0020	.0049	.0321			.00001	.00016
	20	.578	.753	.840	.963	.0584	.140	.220	.511	.00045	.0024	.0057	.0413			.00001	.00024
31	2	.512	.512	.512	.512	.0287	.0287	.0287	.0287	.00012	.00012	.00012	.00012				
	3	.507	.549	.586	.661	.0271	.0342	.0417	.0621	.00009	.00012	.00017	.00044				
	4	.505	.572	.627	.738	.0263	.0388	.0516	.0934	.00007	.00013	.00022	.00065				
	5	.503	.591	.656	.786	.0259	.0421	.0596	.121	.00007	.00015	.00027	.0010				
	6	.502	.606	.678	.819	.0255	.0456	.0669	.147	.00006	.00016	.00032	.0014				
	8	.500	.627	.706	.861	.0251	.0507	.0774	.190	.00006	.00019	.00041	.0024				
	10	.500	.642	.728	.888	.0249	.0548	.0873	.227	.00005	.00021	.00049	.0035				
	12	.499	.654	.745	.906	.0248	.0584	.0957	.259	.00005	.00023	.00057	.0046				
	16	.498	.671	.768	.929	.0246	.0640	.109	.310	.00005	.00026	.00073	.0070				.00001
	20	.498	.686	.786	.944	.0245	.0692	.120	.353	.00005	.00030	.00087	.0096				.00001

*No entry indicates that the value is less than .000005.

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

A. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_{\bar{X}})_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.05$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 0.5$				$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$			
41	2	.392	.392	.392	.392	.0065	.0065	.0065	.0065								
	3	.388	.430	.466	.574	.0060	.0080	.0102	.0167				.00001				
	4	.386	.455	.509	.631	.0058	.0093	.0132	.0278			.00001	.00002				
	5	.385	.473	.536	.688	.0057	.0103	.0158	.0389			.00001	.00004				
	6	.385	.489	.563	.729	.0056	.0114	.0182	.0497			.00001	.00006				
	8	.384	.510	.597	.784	.0055	.0130	.0222	.0705			.00001	.00011				
	10	.383	.526	.621	.820	.0054	.0143	.0256	.0892		.00001	.00001	.00017				
	12	.383	.540	.640	.846	.0053	.0155	.0288	.107		.00001	.00002	.00025				
	16	.382	.559	.668	.879	.0053	.0174	.0340	.138		.00001	.00002	.00042				
	20	.382	.575	.689	.901	.0052	.0195	.0386	.166		.00001	.00003	.00064				
61	2	.218	.218	.218	.218	.00028	.00028	.00028	.00028								
	3	.216	.248	.277	.346	.00025	.00036	.00049	.00094								
	4	.215	.270	.315	.431	.00023	.00043	.00068	.0018								
	5	.214	.285	.343	.494	.00022	.00049	.00086	.0029								
	6	.214	.299	.365	.541	.00021	.00055	.0010	.0041								
	8	.213	.318	.399	.612	.00021	.00066	.0013	.0068								
	10	.213	.333	.424	.660	.00020	.00075	.0016	.0095								
	12	.213	.346	.444	.698	.00020	.00083	.0019	.0126								
	16	.212	.364	.474	.749	.00019	.00097	.0024	.0184								
	20	.212	.379	.498	.785	.00019	.0011	.0029	.0244								
121	2	.0281	.0281	.0281	.0281												
	3	.0278	.0350	.0422	.0627												
	4	.0275	.0405	.0531	.0951												
	5	.0274	.0448	.0622	.125												
	6	.0273	.0488	.0697	.152												
	8	.0272	.0546	.0817	.198												
	10	.0271	.0593	.0927	.235												
	12	.0271	.0636	.102	.269												
	16	.0270	.0699	.116	.322				.00001								
	20	.0270	.0756	.128	.366				.00001								

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

B. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.01$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 2.0$				$\delta' = 2.5$				$\delta' = 3.0$				$\delta' = 3.5$			
5	2	.574	.574	.574	.574	.283	.283	.283	.283	.101	.101	.101	.101	.0305	.0305	.0305	.0305
	3	.458	.521	.557	.653	.193	.238	.266	.352	.0586	.0765	.0886	.131	.0145	.0195	.0230	.0362
	4	.406	.494	.550	.691	.159	.216	.260	.390	.0436	.0649	.0835	.150	.0094	.0148	.0198	.0404
	5	.378	.480	.551	.715	.141	.205	.258	.417	.0361	.0592	.0811	.164	.0071	.0125	.0182	.0441
	6	.359	.477	.554	.734	.130	.202	.260	.439	.0317	.0571	.0808	.177	.0057	.0115	.0175	.0476
	8	.337	.478	.560	.764	.117	.202	.264	.475	.0267	.0557	.0814	.200	.0043	.0105	.0168	.0550
	10	.325	.483	.566	.785	.109	.204	.269	.503	.0240	.0560	.0827	.219	.0036	.0102	.0167	.0617
	12	.317	.487	.574	.802	.105	.207	.275	.527	.0224	.0565	.0850	.237	.0032	.0100	.0169	.0683
	16	.307	.497	.587	.828	.0995	.214	.286	.564	.0204	.0586	.0893	.266	.0027	.0102	.0177	.0802
	20	.301	.508	.599	.846	.0959	.222	.296	.593	.0196	.0612	.0937	.291	.0024	.0106	.0185	.0914
		$\delta' = 1.5$				$\delta' = 2.0$				$\delta' = 2.5$				$\delta' = 3.0$			
7	2	.596	.596	.596	.596	.253	.253	.253	.253	.0654	.0654	.0654	.0654	.0125	.0125	.0125	.0125
	3	.528	.583	.614	.694	.199	.240	.265	.340	.0444	.0573	.0661	.0962	.0068	.0092	.0109	.0171
	4	.496	.575	.628	.744	.177	.232	.275	.395	.0361	.0530	.0681	.120	.0048	.0076	.0103	.0213
	5	.478	.578	.637	.775	.165	.233	.283	.433	.0317	.0524	.0699	.138	.0038	.0070	.0101	.0250
	6	.466	.583	.645	.797	.157	.237	.289	.461	.0290	.0527	.0715	.154	.0032	.0068	.0101	.0283
	8	.451	.593	.660	.828	.147	.244	.302	.508	.0259	.0545	.0759	.182	.0026	.0067	.0104	.0352
	10	.442	.603	.672	.851	.142	.252	.314	.544	.0241	.0565	.0802	.207	.0022	.0068	.0108	.0418
	12	.437	.611	.682	.866	.138	.258	.324	.571	.0230	.0582	.0837	.226	.0020	.0069	.0113	.0476
	16	.430	.625	.699	.889	.134	.270	.341	.615	.0216	.0620	.0906	.261	.0018	.0073	.0125	.0588
	20	.425	.638	.713	.905	.131	.281	.355	.648	.0208	.0659	.0967	.290	.0017	.0078	.0133	.0693
9	2	.399	.399	.399	.399	.102	.102	.102	.102	.0150	.0150	.0150	.0150	.0017	.0017	.0017	.0017
	3	.352	.401	.430	.511	.0806	.100	.114	.156	.0097	.0129	.0153	.0235	.00078	.0011	.0013	.0021
	4	.330	.402	.451	.576	.0711	.100	.122	.196	.0076	.0119	.0159	.0314	.00050	.00083	.0012	.0027
	5	.318	.408	.465	.619	.0659	.102	.129	.227	.0065	.0117	.0164	.0384	.00036	.00074	.0011	.0032
	6	.310	.418	.478	.651	.0625	.105	.135	.252	.0059	.0120	.0172	.0448	.00030	.00070	.0011	.0037
	8	.300	.433	.498	.680	.0587	.112	.145	.295	.0051	.0127	.0187	.0571	.00022	.00068	.0011	.0049
	10	.294	.445	.515	.731	.0562	.117	.154	.330	.0047	.0133	.0201	.0685	.00018	.00069	.0012	.0061
	12	.290	.455	.528	.755	.0548	.122	.162	.357	.0044	.0139	.0214	.0782	.00016	.00070	.0012	.0073
	16	.286	.472	.549	.791	.0530	.130	.175	.402	.0041	.0151	.0239	.0964	.00013	.00073	.0013	.0096
	20	.283	.487	.566	.818	.0519	.138	.186	.440	.0039	.0164	.0261	.114	.00012	.00079	.0015	.0122

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

B. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.01$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$				$\delta' = 2.5$			
11	2	.689	.689	.689	.689	.254	.254	.254	.254	.0399	.0399	.0399	.0399	.0034	.0034	.0034	.0034
	3	.656	.698	.724	.786	.224	.261	.287	.357	.0309	.0394	.0459	.0667	.0020	.0028	.0033	.0054
	4	.639	.708	.746	.833	.210	.270	.308	.423	.0270	.0404	.0506	.0896	.0015	.0025	.0034	.0075
	5	.630	.718	.760	.860	.203	.279	.324	.469	.0248	.0420	.0543	.108	.0013	.0025	.0036	.0096
	6	.623	.726	.771	.879	.197	.286	.336	.504	.0234	.0434	.0576	.125	.0011	.0025	.0037	.0114
	8	.615	.741	.789	.905	.191	.301	.358	.559	.0217	.0468	.0640	.154	.00092	.0026	.0041	.0155
	10	.610	.753	.802	.921	.187	.315	.374	.599	.0208	.0504	.0692	.180	.00082	.0028	.0045	.0194
	12	.607	.764	.811	.932	.185	.326	.387	.629	.0201	.0534	.0737	.200	.00076	.0030	.0048	.0230
	16	.603	.775	.827	.947	.182	.340	.410	.675	.0194	.0574	.0815	.237	.00069	.0032	.0055	.0302
	20	.601	.787	.838	.957	.180	.354	.428	.710	.0190	.0617	.0886	.268	.00065	.0036	.0061	.0373
13	2	.596	.596	.596	.596	.157	.157	.157	.157	.0152	.0152	.0152	.0152	.00076	.00076	.00076	.00076
	3	.567	.615	.641	.710	.138	.167	.184	.239	.0116	.0152	.0177	.0271	.00041	.00058	.00070	.0012
	4	.552	.627	.667	.768	.129	.174	.203	.297	.0098	.0154	.0198	.0382	.00029	.00051	.00071	.0017
	5	.544	.639	.685	.803	.124	.181	.216	.339	.0088	.0161	.0216	.0481	.00023	.00048	.00072	.0022
	6	.539	.649	.699	.827	.122	.188	.227	.373	.0083	.0168	.0232	.0571	.00020	.00048	.00075	.0027
	8	.531	.666	.720	.862	.117	.200	.246	.428	.0076	.0183	.0262	.0743	.00015	.00050	.00083	.0038
	10	.527	.680	.736	.883	.115	.211	.261	.469	.0072	.0198	.0288	.0896	.00013	.00054	.00090	.0049
	12	.524	.691	.748	.898	.113	.220	.273	.502	.0070	.0212	.0311	.103	.00012	.00057	.00098	.0060
	16	.521	.708	.767	.919	.111	.234	.293	.552	.0067	.0234	.0352	.127	.00011	.00063	.0011	.0083
	20	.518	.722	.781	.934	.110	.247	.310	.592	.0065	.0256	.0388	.150	.00010	.00070	.0013	.0107
		$\delta' = 0.5$				$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$			
16	2	.904	.904	.904	.904	.469	.469	.469	.469	.0731	.0731	.0731	.0731	.0034	.0034	.0034	.0034
	3	.896	.915	.925	.948	.445	.491	.520	.595	.0637	.0787	.0896	.123	.0024	.0033	.0040	.0065
	4	.891	.922	.936	.965	.434	.508	.551	.664	.0595	.0844	.100	.163	.0020	.0034	.0045	.0097
	5	.889	.928	.942	.973	.427	.524	.572	.708	.0570	.0899	.111	.195	.0018	.0036	.0049	.0128
	6	.887	.931	.947	.979	.422	.534	.588	.739	.0553	.0936	.118	.222	.0016	.0037	.0053	.0158
	8	.884	.938	.955	.985	.417	.556	.614	.785	.0534	.103	.132	.268	.0015	.0041	.0062	.0220
	10	.883	.944	.959	.989	.414	.573	.633	.815	.0523	.110	.142	.304	.0014	.0045	.0070	.0279
	12	.882	.948	.963	.991	.411	.587	.647	.837	.0515	.117	.151	.333	.0013	.0049	.0076	.0333
	16	.881	.953	.968	.994	.408	.604	.670	.867	.0505	.126	.165	.382	.0012	.0055	.0088	.0439
	20	.880	.957	.971	.996	.407	.620	.688	.888	.0500	.134	.178	.422	.0012	.0061	.0100	.0542

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

B. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.01$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 0.5$				$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$			
21	2	.858	.858	.858	.858	.297	.297	.297	.297	.0185	.0185	.0185	.0185	.00026	.00026	.00026	.00026
	3	.849	.873	.887	.918	.282	.320	.346	.416	.0158	.0205	.0240	.0357	.00016	.00023	.00028	.00050
	4	.844	.885	.903	.943	.274	.341	.378	.491	.0145	.0227	.0284	.0520	.00012	.00023	.00032	.00081
	5	.842	.893	.912	.956	.270	.356	.399	.542	.0138	.0247	.0319	.0665	.00010	.00024	.00036	.0011
	6	.840	.899	.919	.965	.267	.368	.417	.581	.0133	.0263	.0348	.0795	.00009	.00025	.00039	.0015
	8	.837	.909	.931	.975	.264	.389	.445	.640	.0128	.0295	.0402	.104	.00008	.00028	.00046	.0022
	10	.836	.916	.937	.981	.262	.406	.465	.680	.0124	.0323	.0447	.125	.00007	.00031	.00053	.0030
	12	.835	.921	.942	.985	.260	.419	.482	.710	.0122	.0346	.0485	.144	.00007	.00034	.00059	.0039
	16	.834	.928	.949	.989	.258	.438	.507	.754	.0119	.0384	.0550	.176	.00006	.00039	.00071	.0056
	20	.833	.934	.954	.992	.257	.456	.528	.786	.0118	.0422	.0609	.204	.00006	.00044	.00082	.0074
25	2	.817	.817	.817	.817	.199	.199	.199	.199	.0059	.0059	.0059	.0059	.00003	.00003	.00003	.00003
	3	.809	.837	.853	.890	.188	.220	.240	.299	.0049	.0065	.0078	.0123	.00002	.00003	.00003	.00006
	4	.804	.851	.873	.922	.183	.237	.268	.369	.0044	.0073	.0094	.0190	.00001	.00002	.00004	.00010
	5	.801	.862	.885	.940	.180	.251	.288	.420	.0041	.0080	.0107	.0254	.00001	.00002	.00004	.00014
	6	.799	.869	.894	.951	.178	.261	.303	.458	.0040	.0087	.0119	.0316	.00001	.00002	.00004	.00019
	8	.797	.881	.907	.965	.176	.280	.329	.520	.0038	.0099	.0141	.0438	.00001	.00003	.00005	.00031
	10	.796	.889	.915	.973	.174	.295	.349	.564	.0037	.0109	.0160	.0549	.00001	.00003	.00006	.00044
	12	.795	.896	.922	.978	.173	.307	.364	.598	.0036	.0119	.0176	.0649	.00001	.00003	.00007	.00057
	16	.794	.905	.931	.984	.172	.326	.389	.649	.0035	.0135	.0205	.0836		.00004	.00008	.00088
	20	.793	.913	.938	.988	.171	.342	.409	.688	.0034	.0151	.0233	.101		.00005	.00009	.0012
31	2	.754	.754	.754	.754	.103	.103	.103	.103	.00096	.00096	.00096	.00096				
	3	.745	.780	.799	.844	.0974	.118	.131	.173	.00075	.0011	.0013	.0022				
	4	.741	.797	.823	.886	.0947	.130	.151	.226	.00066	.0012	.0016	.0037				
	5	.738	.811	.839	.906	.0931	.140	.165	.267	.00061	.0013	.0019	.0053				
	6	.737	.820	.850	.925	.0919	.148	.178	.302	.00058	.0015	.0021	.0069				.00001
	8	.735	.836	.867	.945	.0908	.162	.197	.358	.00054	.0017	.0026	.0103				.00001
	10	.733	.846	.879	.957	.0898	.172	.213	.402	.00052	.0019	.0030	.0136				.00002
	12	.732	.855	.887	.964	.0894	.182	.225	.436	.00050	.0021	.0034	.0169				.00003
	16	.731	.867	.899	.973	.0887	.196	.246	.490	.00048	.0025	.0041	.0234				.00004
	20	.731	.876	.908	.980	.0883	.209	.263	.533	.00047	.0028	.0047	.0300				.00006

TABLE 2 (continued)

COMPARISON OF TYPE II–III ERROR RATES FOR SINGLE CLASSIFICATION (δ' AT 0.5 INTERVALS)

B. Values of $\beta'_1 = \beta'_L$, $\beta'_2 = (\beta'_S)_{\max}$, $\beta'_3 = (\beta'_X)_{\min}$ and $\beta'_4 = \beta'_W$ for $\alpha = 0.01$

N	m	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4	β'_1	β'_2	β'_3	β'_4
		$\delta' = 0.5$				$\delta' = 1.0$				$\delta' = 1.5$				$\delta' = 2.0$			
41	2	.650	.650	.650	.650	.0309	.0309	.0309	.0309	.00004	.00004	.00004	.00004				
	3	.638	.676	.701	.759	.0292	.0368	.0427	.0609	.00003	.00004	.00005	.00010				
	4	.634	.701	.732	.815	.0283	.0423	.0513	.0872	.00002	.00005	.00007	.00019				
	5	.632	.718	.753	.849	.0277	.0468	.0580	.110	.00002	.00006	.00008	.00030				
	6	.630	.730	.769	.872	.0273	.0503	.0639	.131	.00002	.00006	.00009	.00042				
	8	.628	.751	.791	.903	.0268	.0570	.0739	.168	.00002	.00007	.00012	.00069				
	10	.627	.764	.807	.921	.0265	.0619	.0819	.198	.00002	.00008	.00015	.00099				
	12	.626	.776	.819	.933	.0264	.0666	.0885	.224	.00002	.00009	.00017	.0013				
	16	.625	.792	.836	.949	.0262	.0737	.0995	.267	.00001	.00011	.00021	.0020				
	20	.625	.805	.849	.960	.0260	.0805	.109	.304	.00001	.00013	.00025	.0028				
61	2	.443	.443	.443	.443	.0022	.0022	.0022	.0022								
	3	.437	.482	.506	.575	.0020	.0028	.0033	.0055								
	4	.435	.508	.544	.650	.0019	.0033	.0043	.0091								
	5	.433	.529	.570	.699	.0018	.0038	.0051	.0129								
	6	.432	.544	.590	.736	.0018	.0042	.0059	.0169								
	8	.431	.570	.621	.785	.0017	.0050	.0072	.0246								
	10	.430	.586	.642	.818	.0017	.0056	.0084	.0321								
	12	.429	.602	.658	.840	.0017	.0063	.0094	.0390								
	16	.429	.623	.683	.872	.0017	.0073	.0113	.0521				.00001				
	20	.428	.641	.702	.894	.0017	.0083	.0130	.0650				.00001				
121	2	.0986	.0986	.0986	.0986												
	3	.0974	.118	.130	.170												
	4	.0966	.132	.151	.222												
	5	.0962	.144	.168	.267												
	6	.0960	.153	.182	.305				.00001								
	8	.0956	.170	.204	.364				.00001								
	10	.0954	.181	.221	.409				.00002								
	12	.0953	.191	.234	.442				.00003								
	16	.0951	.207	.256	.500				.00005								
	20	.0950	.221	.274	.545				.00008								

The values in Table 2, like those in Table 1, were computed for the single-classification case, and the values for randomized blocks and other designs are slightly different.

No attempt has been made to specify in the hypothesis under test the values or relative values of population means other than μ_U and μ_L. This does not affect the fixed range tests, and leads to no error in the maximum values of β', but produces an upward bias in minimum values of β' for the multiple range tests. This is true since even if the range of two adjacent sample means is significant, the two will be said to differ significantly only if the ranges of all groups containing them are significant. Wine [1955] has made a power study of multiple range tests in which the relative values of all population means are specified. He tabulated the power for tests on 3 and 4 means. For larger values of m, his approach appears forbiddingly difficult, and the present approach, though not entirely satisfactory, seems to be the only feasible one.

8. *Required Sample Sizes*

If α, β, δ' and m are fixed, the required sample size N for each of the fixed range tests can be determined. For multiple range tests, the required sample size N depends upon p. If α, β, δ' and m are fixed, the minimum required sample size $N_{\min}$ occurs for $p = 2$, and the maximum required sample size $N_{\max}$ occurs for $p = m$.

The code letters for the various tests are used as subscripts on the required sample size N; for example, N_L denotes the required sample size for the LSD test.

For $p = 2$, the critical ranges for the Newman-Keuls test and Duncan's new multiple range test are equal to the LSD; hence $(N_I)_{\min} = (N_S)_{\min} = N_L$. For $p = m$, the critical ranges for the Newman-Keuls test and Tukey's X procedure are equal to the WSD; hence $(N_I)_{\max} = (N_X)_{\max} = N_W$.

Table 3 gives N_L, $(N_S)_{\max}$, $(N_X)_{\min}$, and N_W for $\alpha = 0.05, 0.01$; $\beta = 0.10, 0.05$; $m = 2(1)10(2)20$; $\delta' = 0.5(0.5)3.0$. The required sample sizes N_E for Fisher's test have not been computed, but they are slightly larger than the corresponding N_W.

For small values of δ', the required sample size N is approximately inversely proportional to δ'^2. Hence, for $\delta' < 1$, the approximate value of $N\delta'^2$ may be read from the column headed $\delta' = 1.0$ in Table 3.

The values in Table 3 have been computed for the single-classification case. To obtain sample sizes which will ensure that α and β do not exceed the prescribed values for randomized blocks, it can be shown that it is sufficient (though not always necessary) to add 2 to the sample sizes for $m = 2$ and to add 1 for $m \geq 3$.

TABLE 3

SAMPLE SIZE N FOR PRESCRIBED α AND β [$\delta' = 0.5\ (0.5)\ 3.0$]
(SINGLE CLASSIFICATION)

A. $\alpha = 0.05$, $\beta = 0.10$ ($P = 0.90$)

	m	$\delta' = 0.5$	$\delta' = 1.0$	$\delta' = 1.5$	$\delta' = 2.0$	$\delta' = 2.5$	$\delta' = 3.0$
	2	86	23	11	7	5	4
	3	85	22	10	7	5	4
N_L	4	85	22	10	6	4	4
	5–20	85	22	10	6	4	3
	2	86	23	11	7	5	4
	3	91	24	11	7	5	4
	4	94	24	11	7	5	4
	5	97	25	12	7	5	4
	6	99	25	12	7	5	4
	7	101	26	12	7	5	4
	8	102	26	12	7	5	4
$(N_S)_{\max}$	9	104	26	12	7	5	4
	10	105	27	12	7	5	4
	12	107	27	12	7	5	4
	14	108	28	13	7	5	4
	16	110	28	13	7	5	4
	18	111	28	13	7	5	4
	20	112	28	13	8	5	4
	2	86	23	11	7	5	4
	3	96	25	12	7	5	4
	4	102	26	12	7	5	4
	5	106	27	13	8	5	4
	6	110	28	13	8	5	4
	7	113	29	13	8	5	4
	8	115	29	14	8	5	4
$(N_X)_{\min}$	9	117	30	14	8	6	4
	10	119	30	14	8	6	4
	12	122	31	14	8	6	4
	14	125	32	15	9	6	4
	16	127	32	15	9	6	4
	18	129	33	15	9	6	4
	20	131	33	15	9	6	4
	2	86	23	11	7	5	4
	3	106	28	13	8	6	4
	4	120	31	14	9	6	5
	5	130	33	15	9	6	5
N_W	6	138	35	16	10	7	5
	7	144	37	17	10	7	5
	8	150	38	17	10	7	5
	9	155	39	18	11	7	5
	10	159	40	18	11	7	5
	12	166	42	19	11	7	5
	14	173	44	20	12	8	6
	16	178	45	20	12	8	6
	18	183	46	21	12	8	6
	20	187	47	21	12	8	6

TABLE 3 (continued)

B. $\alpha = 0.05$, $\beta = 0.05$ ($P = 0.95$)

	m	$\delta' = 0.5$	$\delta' = 1.0$	$\delta' = 1.5$	$\delta' = 2.0$	$\delta' = 2.5$	$\delta' = 3.0$
	2	106	28	13	8	6	5
	3	105	27	13	8	6	4
N_L	4	105	27	13	8	5	4
	5	105	27	13	7	5	4
	6-20	105	27	12	7	5	4
	2	106	28	13	8	6	5
	3	111	29	13	8	6	4
	4	115	30	14	8	6	4
	5	119	30	14	8	6	4
	6	121	31	14	8	6	4
	7	122	31	14	8	6	4
	8	124	32	14	8	6	4
$(N_S)_{\max}$	9	126	32	15	9	6	4
	10	127	32	15	9	6	4
	12	129	33	15	9	6	4
	14	131	33	15	9	6	4
	16	132	33	15	9	6	4
	18	134	34	15	9	6	4
	20	135	34	15	9	6	4
	2	106	28	13	8	6	5
	3	116	30	14	9	6	5
	4	123	32	15	9	6	5
	5	128	33	15	9	6	5
	6	132	34	16	9	6	5
	7	135	35	16	9	6	5
	8	138	35	16	9	6	5
$(N_X)_{\min}$	9	140	36	16	10	6	5
	10	142	36	17	10	7	5
	12	146	37	17	10	7	5
	14	149	38	17	10	7	5
	16	151	38	17	10	7	5
	18	153	39	18	10	7	5
	20	155	39	18	10	7	5
	2	106	28	13	8	6	5
	3	129	33	16	9	7	5
	4	143	37	17	10	7	5
	5	154	39	18	11	7	5
N_W	6	163	41	19	11	8	6
	7	170	43	20	12	8	6
	8	176	45	20	12	8	6
	9	181	46	21	12	8	6
	10	186	47	21	12	8	6
	12	194	49	22	13	9	6
	14	201	51	23	13	9	6
	16	206	52	24	14	9	7
	18	211	53	24	14	9	7
	20	216	55	25	14	9	7

TABLE 3 (continued)

C. $\alpha = 0.01$, $\beta = 0.10$ ($P = 0.90$)

	m	$\delta' = 0.5$	$\delta' = 1.0$	$\delta' = 1.5$	$\delta' = 2.0$	$\delta' = 2.5$	$\delta' = 3.0$
	2	121	32	15	10	7	6
	3	121	31	15	9	6	6
	4	120	31	15	9	6	5
N_L	5	120	31	14	9	6	5
	6–13	120	31	14	8	6	5
	14–15	120	30	14	8	6	5
	16–20	120	30	14	8	5	5
	2	121	32	15	10	7	6
	3	127	33	16	10	7	5
	4	132	34	16	10	7	5
	5	135	35	16	10	7	5
	6	138	35	16	10	7	5
	7	140	36	16	10	7	5
$(N_S)_{\max}$	8	143	36	17	10	7	5
	9	144	37	17	10	7	5
	10	145	37	17	10	7	5
	12	148	38	17	10	7	5
	14	151	38	17	10	7	5
	16	152	39	18	10	7	5
	18	154	39	18	10	7	5
	20	155	39	18	10	7	5
	2	121	32	15	10	7	6
	3	131	34	16	10	7	5
	4	138	35	17	10	7	5
	5	142	36	17	10	7	5
	6	146	37	17	10	7	5
	7	149	38	17	10	7	5
$(N_X)_{\min}$	8	151	39	18	10	7	5
	9	153	39	18	11	7	5
	10	155	40	18	11	7	5
	12	158	40	18	11	7	5
	14	161	41	19	11	7	5
	16	163	41	19	11	7	5
	18	165	42	19	11	7	5
	20	167	42	19	11	7	5
	2	121	32	15	10	7	6
	3	143	37	17	11	7	6
	4	156	40	19	11	8	6
	5	166	43	20	12	8	6
	6	174	44	20	12	8	6
	7	180	46	21	12	8	6
N_W	8	186	47	22	13	9	6
	9	191	49	22	13	9	7
	10	195	50	23	13	9	7
	12	203	51	23	14	9	7
	14	209	53	24	14	9	7
	16	214	54	25	14	9	7
	18	219	55	25	14	10	7
	20	224	56	26	15	10	7

TABLE 3 (continued)

D. $\alpha = 0.01, \beta = 0.05$ ($P = 0.95$)

	m	$\delta' = 0.5$	$\delta' = 1.0$	$\delta' = 1.5$	$\delta' = 2.0$	$\delta' = 2.5$	$\delta' = 3.0$
	2	145	38	18	11	8	6
	3	144	37	17	11	7	6
	4–7	144	37	17	10	7	5
N_L	8	143	37	17	10	7	5
	9–12	143	36	17	10	7	5
	13–19	143	36	17	10	6	5
	20	143	36	16	10	6	5
	2	145	38	18	11	8	6
	3	151	39	18	11	8	6
	4	156	40	19	11	8	6
	5	160	41	19	11	8	6
	6	163	42	19	11	8	6
	7	166	42	19	11	8	6
$(N_S)_{\max}$	8	168	43	19	11	8	6
	9	170	43	20	11	8	6
	10	171	43	20	12	8	6
	12	174	44	20	12	8	6
	14	177	45	20	12	8	6
	16	179	45	20	12	8	6
	18	181	46	21	12	8	6
	20	182	46	21	12	8	6
	2	145	38	18	11	8	6
	3	156	40	19	11	8	6
	4	163	42	19	12	8	6
	5	168	43	20	12	8	6
	6	172	44	20	12	8	6
	7	175	44	20	12	8	6
$(N_X)_{\min}$	8	177	45	21	12	8	6
	9	180	46	21	12	8	6
	10	182	46	21	12	8	6
	12	185	47	21	12	8	6
	14	188	48	22	13	8	6
	16	191	48	22	13	8	6
	18	193	49	22	13	8	6
	20	195	49	22	13	9	6
	2	145	38	18	11	8	6
	3	168	43	20	12	8	7
	4	183	47	22	13	9	7
	5	193	49	23	13	9	7
N_W	6	202	51	24	14	9	7
	7	209	53	24	14	10	7
	8	215	55	25	15	10	7
	9	220	56	25	15	10	7
	10	225	57	26	15	10	7
	12	233	59	27	15	10	7
	14	240	61	27	16	10	8
	16	245	62	28	16	11	8
	18	250	63	29	16	11	8
	20	255	64	29	17	11	8

9. *Computation of the Tables*

Let the critical range MSD of a range test on p out of m ordered means with significance level α_M be represented by

$$\text{MSD} = Q(\alpha_M , p, m, \nu)s_{\bar{x}} , \tag{18}$$

where M may be I, X, S, or W and Q (α_M , p, m, ν) is the significant studentized range at the α_M level for a test on p out of m ordered means with ν degrees of freedom for s. The equivalent LSD test has significance level α_L such that

$$t(\alpha_L , \nu) = Q(\alpha_M , p, m, \nu)/\sqrt{2}. \tag{19}$$

It should be noted that (19) is a generalization of (8). While (8) holds only for Tukey's studentized range test, (19) holds for that test and also for the multiple range tests. The part of Table 1 concerned with α_L corresponding to $\alpha_E = 0.05, 0.01$ is computed directly from (9). The remainder of Table 1 is computed from (19). One knows $Q(\alpha_M , p, m, \nu)$ and must find α_L by interpolation from the probability integral of the t-distribution, which has been tabulated by Hartley and Pearson [1950]. The method of interpolation is linear harmonic ν-wise and three-point Lagrangian t-wise. For $\nu > 20$, $t > 4$, the Hartley-Pearson tables cannot be used. With $t > 4$, tables of the incomplete Beta-function edited by Karl Pearson [1934] were used for $\nu = 21$ and for even values of ν up to and including 100. The probability that t exceeds t_0 is given by

$$P(t > t_0) = I_x(1/2, \nu/2), \tag{20}$$

where

$$x = \nu/(\nu + t_0^2). \tag{21}$$

In one case ($\nu = 27$, $t > 4$), neither of the above methods is applicable, but there a method proposed by Fettis [1954] can be used.

The combined Type II and Type III error rate, β'_M , is given by the relation

$$\beta'_M = P\{t < [Q(\alpha_M , p, m, \nu) - \delta'\sqrt{N}]/\sqrt{2}\}. \tag{22}$$

In computing Table 2, one knows α_M , m, N, p and δ', and hence also the value below which t must lie. The corresponding probability is found by use of the Hartley and Pearson [1950] tables of the probability integral of the t-distribution or one of the other methods mentioned above. It would perhaps be better from a theoretical point of view to fix δ rather than δ', but then in computing tables it would be necessary to use the non-central t-distribution. Available tables of this distri-

bution, including those published recently by Resnikoff and Lieberman [1957], are such that this does not appear feasible.

Table 3 is also computed from equation (22) by one of the above methods. Here α_M , β'_M , m, p and δ' are known, and N is to be determined. Since only integral values of N are possible, the value of N found by solving equation (22) must be rounded to the next higher integer.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to D. B. Duncan, H. O. Hartley, and J. W. Tukey, who have made available certain of their unpublished papers and memoranda, and have offered a number of helpful suggestions.

REFERENCES

Beyer, W. H. [1953]. *Certain Percentage Points of the Distribution of the Studentized Range of Large Samples*, Virginia Polytechnic Institute Technical Report No. 4 (Sponsored by Office of Ordnance Research, U.S. Army).

Duncan, D. B. [1953]. *Biometrics 9*, Abstract 220.

Duncan, D. B. [1955]. *Biometrics 11*, 1.

Federer, W. T. [1955]. *Experimental Designs*, The Macmillan Company, New York.

Fettis, H. E. [1954]. *Journal of Mathematics and Physics 33*, 283.

Fisher, R. A. [1935]. *The Design of Experiments*. Oliver and Boyd, Ltd., Edinburgh.

Hartley, H. O. and Pearson, E. S. [1950]. *Biometrika 37*, 168.

Keuls, M. [1952]. *Euphytica 1*, 112.

Newman, D. [1939]. *Biometrika 31*, 20.

Pearson, E. S. and Hartley, H. O. [1954]. *Biometrika Tables for Statisticians*. Cambridge University Press, Cambridge.

Pearson, K. [1934]. *Table of the Incomplete Beta-Function*. Biometrika Office, London.

Resnikoff, G. J. and Lieberman, G. J. [1957]. *Tables of the Non-central t-Distribution*. Stanford University Technical Report No. 32. (Sponsored by Office of Naval Research).

"Student" [1908]. *Biometrika 6*, 1.

Tukey J. W. [1952]. *Allowances for Various Types of Error Rates*, unpublished invited address presented at Blacksburg meeting of Institute of Mathematical Statistics.

Tukey J. W. [1953]. *The Problem of Multiple Comparisons*, unpublished memorandum in private circulation.

Wine, R. L. [1955]. *A Power Study of Multiple Range and Multiple F Tests*. Virginia Polytechnic Institute Technical Report No. 12 (Sponsored by Office of Ordnance Research, U. S. Army).

QUERIES AND NOTES

GEORGE W. SNEDECOR, *Editor*

127 NOTE: SOME SIMPLE METHODS OF COMPUTING PARAMETERS IN THE ANALYSIS OF VARIANCE

WILLIAM G. MADOW
Stanford Research Institute, Menlo Park, California, U. S. A.

1. *Introduction*

The analysis of variance and F-distributions depend on two parameters, the degrees of freedom, f, of the numerator and denominator mean squares.

Tang's distribution depends not only on the degrees of freedom but also on a parameter, λ, that determines the power of the analysis of variance test.

The purpose of this note is to present some methods of computing both f and λ.

2. *Computing f.*

We begin by recalling that the sums of squares that occur in the analysis of variance are quadratic forms, and that the number of degrees of freedom of the sum of squares is the rank of the quadratic form.

For example, upon expanding the square, we see that

$$S^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i,j=1}^{n} c_{ij} y_i y_j = q$$

where $c_{ii} = 1 - 1/n$ and $c_{ij} = c_{ji} = -1/n$; and $\bar{y}$ is the arithmetic mean of $y_1, \cdots, y_n$. The number of degrees of freedom of S^2 is $n - 1$, which is also the rank of q.

Now the rank of a quadratic form does not depend on the statistical assumptions that may be made. Consequently, we shall assume, in order to compute f, that the underlying random variables are normally and independently distributed with 0 mean and constant variance σ^2.

We shall also assume that q/σ^2 is known to have the χ^2-distribution. Various methods of obtaining analysis of variance ratios guarantee

this; for example, the likelihood ratio test of the general linear hypothesis.

Furthermore, if q/σ^2 has the χ^2-distribution then

$$E\left(\frac{q}{\sigma^2}\right) = f$$

where f is the number of degrees of freedom or rank of q. Thus one way to compute the degrees of freedom is to calculate the expected value of q.

Rule 1. Compute the expected value of q under the hypotheses stated above. Then

$$E(q) = f\sigma^2$$

where f is the number of degrees of freedom.

For example, $E\sum_{i=1}^{n}(y_i - \bar{y})^2 = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$ and hence $f = n - 1$.

But rule 1 can be simplified. Since

$$q = \sum_{i,j=1}^{n} c_{ij}y_iy_j \ ,$$

it follows that if the uncorrelated random variables $y_1 \ , \cdots \ , y_n$ have 0 means and constant variance σ^2, then

$$E(q) = \sigma^2 \sum_{i=1}^{n} c_{ii}$$

and hence $f = \sum_{i=1}^{n} c_{ii}$, (the trace of q).

Rule 2. Determine c_{ii} by putting $y_i = 1$ and $y_j = 0, j \neq i$ in q. (Take advantage of symmetry in doing so.) Thus, in

$$\sum_{i=1}^{n}(y_i - \bar{y})^2, c_{11} = \left(1 - \frac{1}{n}\right)^2 + (n-1)\frac{1}{n^2} = \frac{n-1}{n} = c_{ii} \ ,$$

$$i = 1, \cdots, n.$$

For rule 2 one need not evaluate any expected values. But for many quadratic forms the work of computing f can be further reduced by using two simple theorems on expected values.

(1) If $y_1 \ , \cdots \ , y_m$ are uncorrelated, and have common mean and variance $\sigma^2_{y_i}$, then

$$\sigma^2_{y_i - \bar{y}} = E(y_i - \bar{y})^2 = \frac{n-1}{n}Ey_i^2 = \frac{n-1}{n}\sigma^2_{y_i} \ .$$

(2) Under the same assumptions as (1) we have

$$E \sum_{i=1}^{n} (y_i - \bar{y})^2 = (n-1)\sigma_y^2 .$$

It is easy to prove either (1) or (2); the other is an immediate consequence.

Now let us illustrate the method. Let a dot, $(\cdot)$, indicate the mean for the subscript and let

$$q = \sum_{i=1}^{r} \sum_{j=1}^{s} (x_{ij} - x_{i.} - x_{.j} + x_{..})^2 .$$

Since the expected value of a sum of random variables is the sum of their expected values, we need to evaluate

$$A_{ij} = E(x_{ij} - x_{i.} - x_{.j} + x_{..})^2 .$$

Now

$$A_{ij} = \frac{(r-1)}{r} E(x_{ij} - x_{.j})^2 \quad \text{by (1) since} \quad x_{.j} - x_{..} = \frac{1}{r} \sum_{i=1} (x_{ij} - x_{i.})$$

$$= \frac{(r-1)(s-1)}{rs} \sigma_{xij}^2 , \quad \text{by (1) since} \quad x_{i.} = \frac{1}{s} \sum_{j=1}^{s} x_{ij} .$$

Hence

$$E(q) = (r-1)(s-1)\sigma_{xij}^2$$

and

$$f = (r-1)(s-1).$$

The approach embodied in the above example can be used on many sums of squares especially including those occurring in factorial designs.

Rule 3. If q is a quadratic form which can be written

$$q = \sum_{i=1}^{r} (y_i - \bar{y})^2$$

where $y_1 , \cdots , y_r$ are uncorrelated, have common means, and common variance, σ_{yi}^2 , then

$$E(q) = (r-1)\sigma_{yi}^2$$

and, if $y_i = \xi_{ij} - \xi_{i.}$, where $\xi_{i.}$ is the mean of s uncorrelated random variables $\xi_{i1} , \xi_{i2} , \cdots , \xi_{is}$, which have 0 means and common variance $\sigma^2\xi_{ij}$, then

$$E(q) = (r-1)(s-1)\sigma^2\xi_{ij}$$

and, if the process can be continued on ξ_{ij} ,

$$E(q) = (r - 1)(s - 1)(t - 1)\sigma^2_{n_{ijk}}$$

and so on.

3. *Computing* λ. It is easy to verify that, if $E(y_i) = a_i$, then

$$E(q) = \sigma^2 f + \sum_{i,j=1}^{n} c_{ij}a_i a_j \ .$$

and it has earlier [1] been shown that

$$\lambda = \sum_{i,j=1}^{n} c_{ij}a_i a_j \ .$$

Thus, λ is the same quadratic form in a_1 , $\cdots$, a_n that q is in y_1 , $\cdots$, y_n .

Rule 4. If $E(y_i) = a_i$ instead of 0, λ is determined by replacing y_i by a_i in S^2 where S^2 is the sum of squares that becomes q when the squares are expanded.

Suppose, for example, that $E(y_i) = a$. Then if

$$S^2 = \sum_{i=1}^{n} (y_1 - \bar{y})^2,$$

we have

$$\lambda = \sum_{i=1}^{n} (a - a)^2 = 0.$$

If

$$S^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} (x_{ij} - x_{i.} - x_{.j} + x_{..})^2$$

and $E(x_{ij}) = a_i + b_j$, then

$$\lambda = \sum_{i=1}^{r} \sum_{j=1}^{s} [(a_i + b_j) - (a_i + \bar{b}) - (\bar{a} + b_j) + (\bar{a} + \bar{b})]^2 = 0.$$

But if $E(x_{ij}) = a_{ij}$, then

$$\lambda = \sum_{i=1}^{r} \sum_{j=1}^{s} (a_{ij} - a_{i.} - a_{.j} + a_{..})^2.$$

REFERENCE

[1] W. G. Madow [1940]. The distribution of quadratic forms in non-central normal random variables. *Ann. Math. Stat.*, *11*: 100–104.

128 NOTE: A ROUTINE FOR COMPUTING THE DEGREES OF FREEDOM IN ANALYSIS OF VARIANCE

BERTIL MATÉRN
Forest Research Institute, Stockholm 51, Sweden

This note describes a simple method of automatically computing the degrees of freedom along with the sums of squares in an analysis of variance. The procedure seems to be useful in routine computational work. It may be known to many statisticians, but seems not to be described in any textbook.

A few, rather obvious, remarks will suffice to justify the procedure. We consider a number of observations y_1 , y_2 , $\cdots$. We assume that we can write $y_i = m_i + e_i$, where the "errors" $\{e_i\}$ are normally and independently distributed around zero with constant variance V. By z_1 , z_2 , $\cdots$, we denote linear forms in the y's with error variance V. Thus every z is of the form $\sum a_i y_i$ with $\sum a_i^2 = 1$. In standard applications we start by computing certain sums P_1 , P_2 , $\cdots$, where every P is a sum of squares of some z's. We then form sums of squares attributed to different sources of variation by calculating expressions of the type $Q = \sum g_i P_i$, where the g's are constants (positive, negative, or zero). In all regular cases Q/V has a chi-square distribution under the additional hypothesis that the source of variation in question does not affect the m's. The number of degrees of freedom associated with Q then equals the expectation of Q/V. From the definitions of $\{P_i\}$ and $\{z_i\}$, it is seen that this expectation is $q = \sum g_i p_i$, where p_i is

TABLE 1

EXPERIMENTAL OBSERVATIONS (FICTITIOUS)

Treatment	Block I	Block II	Sum
A_0B_0	2.7	2.9	5.6
A_0B_1	2.0	3.6	5.6
A_1B_0	3.1	3.3	6.4
A_1B_1	3.8	2.7	6.5
A_2B_0	2.5	4.3	6.8
A_2B_1	4.3	4.0	8.3
Sum	18.4	20.8	39.2

Treatment totals:
A_0 : 11.2 A_1 : 12.9 A_2 : 15.1 B_0 : 18.8 B_1 : 20.4

the number of squared z's in P_i . Thus, Q and q are obtained by the same linear operation carried out on the P's and p's, respectively. Therefore, Q and q can conveniently be computed parallelly.

TABLE 2

CALCULATION OF $\{P_i\}$ AND $\{p_i\}$

Grouping	Sum of squares	Number of squares
(1) Individual observations	$P_1 = 2.7^2 + \cdots + 4.0^2 = 134.1200$	$p_1 = 12$
(2) AB	$P_2 = \frac{5.6^2 + \cdots + 8.3^2}{2} = 130.5300$	$p_2 = 6$
(3) A	$P_3 = \frac{11.2^2 + 12.9^2 + 15.1^2}{4} = 129.9650$	$p_3 = 3$
(4) B	$P_4 = \frac{18.8^2 + 20.4^2}{6} = 128.2667$	$p_4 = 2$
(5) Blocks	$P_5 = \frac{18.4^2 + 20.8^2}{6} = 128.5333$	$p_5 = 2$
(6) Total	$P_6 = \frac{39.2^2}{12} = 128.0533$	$p_6 = 1$

TABLE 3

ANALYSIS OF VARIANCE

Computation, from rows of table 2	Source of variation	s.s.	d.f.	m.s.
(5) − (6)	Blocks	0.4800	1	0.4800
(3) − (6)	A	1.9117	2	0.9558
(4) − (6)	B	0.2134	1	0.2134
(2) − (3) − (4) + (6)	AB	0.3516	2	0.1758
(1) − (2) − (5) + (6)	Error	3.1100	5	0.6220
(1) − (6)	Total	6.0667	11	

Example (interaction AB):

$Q = P_2 - P_3 - P_4 + P_6 = 0.3516$

$q = p_2 - p_3 - p_4 + p_6 = 2$

To give an example, we consider a 3×2 factorial experiment in 2 blocks. The factor A has the three levels A_0, A_1, and A_2. The factor B has the two levels B_0, and B_1. Table 1 shows the data of the experiment. The computation of the sums P is given, together with the numbers p, in Table 2. Finally, the sums of squares, and degrees of freedom, as derived by linear operations on P and p, respectively, are shown in Table 3. The example is only intended as an illustration of the computational technique. The procedure is, of course, more useful in complicated cases.

CORRECTIONS TO PAPERS

CHIN LONG CHIANG. **An Application of Stochastic Processes to Experimental Studies on Flour Beetles.** *Biometrics, 13(1)*, 79–97, 1957.

On page 88, in line 20, $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_s)$ should be read as $\tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2, \cdots, \tilde{\theta}_s)$; in line 29, $\zeta_i(\theta)$ and in line 30, $\zeta_i(\hat{\theta})$ should both be read as $\zeta_i(\tilde{\theta})$. On page 90, in equation (48), d_{ui} should be read as d_{ui}.

F. J. ANSCOMBE. **Fixed-Sample-Size Analysis of Sequential Observations.** *Biometrics, 10(1)*, 89–100, 1954.

The two consecutive sentences half-way down p. 100, in which the names of Fisher and Wald are mentioned, should be deleted. It is true that in some instances a distribution obtained by Fisher's fiducial argument, or a decision function obtained by Wald's minimax rule, could have been derived from Bayes' theorem. But I had not appreciated, when writing this paper, that the "prior" distribution that must be inserted in Bayes' theorem in order to obtain Fisher's or Wald's result, may itself depend on the sampling rule. The matter is discussed in a paper, "Dependence on the fiducial argument on the sampling rule," to appear in *Biometrika, 44* (1957).

H. R. THOMPSON. **Extensions to Missing Plot Techniques.** *Biometrics, 12(3)*, 241–244, 1956.

I am grateful to George Hirschhorn for pointing out an arithmetical error in one element of the matrix on page 244 of my paper referenced above. The leading element should be 43470 and not 44345. All entries in the matrix were obtained from the solution for $p = 3$ and divided by 4 to reduce the number of digits. The missing values in Table I are correct, however, as they were calculated using the original correct matrix.

APPRECIATION

When Dr. J. W. Hopkins assumed responsibility for *Biometrics*, in January 1956, the Society looked forward with anticipation to a long and successful term with the journal under his editorship.

Members will learn with regret that John Hopkins later had the misfortune to become ill, and was forced to relinquish office in March of this year, after having fought a losing battle for many months with recurrences of the illness. He felt very keenly his inability to carry on with the job he had undertaken for the Society and the journal, and it was characteristic of the man that he did not wish to retain a responsibility which he could not effectively discharge.

We are greatly indebted to John Hopkins for the sterling service he has given, not only for his contribution as Editor, when the major part of his work was carried out under extreme difficulties, but also earlier as Chairman of a former Finance Committee, as a Councillor, and as an Associate Editor. I take this opportunity of extending to him, on behalf of the Society's membership, our warmest thanks, and best wishes for a rapid recovery.

Thanks are also due to Mrs. Florence H. Suddon, for the unselfish and competent assistance she gave the Editor, especially during the periods of his enforced absences.

We gratefully acknowledge, too, the very generous contribution made by the National Research Council of Canada, in providing office accommodation and the numerous facilities necessary for the efficient conduct of an Editor's duties.

E. A. Cornish
President

ABSTRACTS

Papers presented at the meeting of the Biometric Society (W.N.A.R.) with the American Institute of Biological Sciences at Stanford University, Stanford, California, August 27 and 28, 1957.

450 ROSEDITH SITGREAVES (Teachers College, Columbia University). **Discriminating Analysis in a Non-Classical Case.**

In the classical theory of discriminatory analysis, a linear function (i.e., Fisher's discriminant function) of a set of measurements for a given individual is used in classifying the individual as belonging to one of two populations. This function has been shown to possess some optimal properties if the measurements are continuous, and are normally distributed in each population with the same covariance matrix. Some recent work of Bahadur has been concerned with the problem of classification when the measured variables are all dichotomous, and he has developed some approximations to the optimal classification statistic for this case. In the present paper, a probability model is constructed for an experimental situation in psychology involving "aggressive" and "non-aggressive" children, and Bahadur's results are used to determine an appropriate classification index. This index is evaluated in terms of the probabilities of misclassification. Approximations to the probability distributions and the resulting classification index are also considered.

451 ALBERT C. WALKER (International Minerals and Chemical Corporation, Woodland, California). **Canonical Vectors among Sugar Beet Varieties.**

A canonical analysis of the results of an experiment on thirteen varieties of beets is presented. Nine of these had been chosen because of high sucrose content. Highly significant differences were present both in three commercial characteristics, root weight, sucrose, and glutamate content, and in the three auxiliary measurements of top weight, sodium, and amino acids. Although there was considerable correlation between some of these variables, no reduction in the number of dimensions necessary to depict the varietal differences could be

obtained. The canonical vectors obtained are interpreted as the best available measures of the underlying genetic differences. The order of importance, for this experiment, was found to be (1) vigor, or all-over utilization of energy, (2) distribution of energy between growth and storage in the form of sucrose or glutamate, (3) nitrogen foraging ability, (4) distribution of growth between root and top, and (5) a curious positive correlation between sodium and sucrose, in two varieties known to be carrying a gene causing albinism.

452 WILLIAM F. ROYCE (U. S. Fish and Wildlife Service, Juneau). **Statistical Comparisons of Morphological Data.**

The methods of statistical comparison of morphological data are examined with special regard to techniques suitable for racial studies of tunas. Comparison of samples by means of a test of significance is discarded because it leads to a trivial conclusion. In its place is suggested and defined a measure of overlap of frequency distributions.

The computation of overlap for single counted characters is extended to single measured characters through the use of regression analysis and to multiple characters either counted or measured or both.

The general concept of overlap is related to the proportion of one sample which might belong to another specified sample of equal size. This is considered to be a maximum for any intermingling which might have occurred. It is pointed out that such a maximum is useful when considered in conjunction with tagging returns.

453 HARLEY B. MESSINGER (University of California, Berkeley). **Clinical Trials in Breast Cancer, Statistical Problems.**

The commonest measures of treatment effectiveness in cancer are based on survival time after diagnosis or treatment. In the case of breast cancer, there is a strong dependence of survival time upon the age of the patient. Actuarial analysis of the excellent data of the Connecticut Cancer Registry reveals that this age effect can be removed by dividing the survival percentage at each point in time after diagnosis or treatment by the percentage that would be expected to survive to that point in time among a group of people of the same age, sex, and race from the general population.

This fact can be utilized, for example, in the Bross sequential method of treatment comparison to avoid the need for matching pairs of patients by age (see Billewicz' article in Biometrics 12:283, 1956). In each unmatched pair, one would simply observe which patient lived for the greater fraction of her expected survival time.

The common fractions obtained with above correction from the various age groups by weighted averaging increase with the log of time like a cumulative normal distribution function. The lognormal distribution has been fitted before to uncorrected survival curves, but this technique has worked well only in cases where the mortality from cancer was so high that normal mortality was negligible, e.g. lung cancer.

Further, there are indications that the lognormal distribution describes well the distribution of intervals from treatment until the appearance of metastases as well as certain other similar classes of intervals which might have utility as measures of treatment effectiveness. Thus, it may be possible to combine other measures of effectiveness with survival time by the use of multivariate normal techniques to permit much earlier decisions to be made regarding the relative merit of two or more treatments.

MOHAMED S. AHMED (University of California, Berkeley).

454 **Test that the Probability of Disease is Independent of the Total Length of Time Previously Suffered from the Disease.**

Consider an individual who has suffered from a disease for a total length of time τ. Let $P(t; \tau)$ be the probability that, if he contracts the disease again, he remains sick for at least time t. Assume that $P'(t; \tau)\,|_{t=0} = -q(\tau)$. The distribution of the total time T spent in sickness in n successive returns is obtained. Let $q(\tau) = a + 2b\tau$, with a and b non-negative and a known. We test the hypothesis of independence, $b = 0$. The test criterion is to reject the hypothesis of independence whenever $T < k(n, \alpha)/a$, where $k(n, \alpha)$ is adjusted in accordance with the desired level of significance α. The test is uniformly most powerful for the restricted class of alternatives $b \geq a^2 k(n, \alpha)$. A formula for the power function is given. Similar results can be obtained for other situations for example, $q(\tau) = a + 2b\tau + 3c\tau^2$. The problem is more complicated when more than one individual is considered. A method for the estimation of parameters is given and asymptotically similar tests are worked out for this case. Other applications include (*i*) accident-free flying experience of pilots and (*ii*) tunnelling of flour beetles.

W. F. TAYLOR (School of Aviation Medicine, Randolph Field, Texas).

455 **Some Monte Carlo Methods Applied to an Epidemic of Acute Respiratory Disease.**

A probabilistic model and its deterministic counterpart are developed to describe the course of acute respiratory disease among recruits at

the former Air Force reception center at Sampson Air Force Base, New York. Four parameters are estimated and the model is evaluated by comparing observed epidemics which occurred among 117 groups of recruits with 415 analogous Monte Carlo epidemics resulting from using the model. Computations were performed with the aid of a CRC 102A electronic computer.

The following assumptions were made in building the model:

1. Sources of infection exist both from within and without the group.

2. The risk of a susceptible man becoming infected from within the group in a small increment of time, Δt, is proportional to the number of infectious men, I, within the group. Risk from within $= \lambda I$.

3. The risk of infection from without the group is constant. Risk from without $= \mu$.

4. Once recovered from infection, a man has partial immunity, and the constants λ and μ in 2 and 3 above are both altered by multiplying them by the same factor. Risk from within $= \alpha\lambda I$. Risk from without $= \alpha\mu$.

5. There exists an unknown proportion, β, of men who become infected but are not detected. These men remain in the group and provide the pool of infection from within.

456 J. GURLAND AND S. K. KATTI (Iowa State College, Ames, Iowa). **Generalized Families of Contagious Distributions.**

The notion of compound and generalized distribution is utilized to obtain the probability generating functions of certain families of contagious distributions. These families contain the family of Beall and Rescia (Biometrics, 1953) as a special case. Properties of the confluent hypergeometric function are applied to obtain formulae for computing probabilities. Some limiting cases of the families of distributions are also considered when certain of the parameters are allowed to become infinite. In particular, it is shown that the Neyman Type A and the generalized Polya-Aeppli distributions emerge as the limiting form in some cases. Different methods of estimation, including the method of moments and estimation by frequencies, are also considered in a preliminary investigation of some data.

457 DOUGLAS G. CHAPMAN (University of Washington, Seattle, Washington). **Problems of Estimation of Wildlife Mortality Rates.**

The estimation of mortality rates of wild animals in general has been based on one of two methods—recoveries of tagged or marked members

of the population and sampling of the population to determine its age structure at one or more time periods. The sources of bias in these methods are discussed. Several different formulae are reviewed and a new formula given which may be useful for those fish and animals where the largest cause of mortality is predation by man. The effect of the various sources of bias on the formulae is noted. Consideration is also given to the determination of estimates of the variances of the estimates; it will be usually preferable to find internal rather than external estimates of such variances.

B. O. BERGH (University of California, Riverside, California.)
458 **Dominance Estimates from Tomato Populations Segregating for Locule Number.**

Tests of scale adequacy indicated the desirability of a logarithmic transformation of the data. Despite some evidence that this transformation was not strong enough, it was statistically acceptable according to the tests used.

Estimates of the average absolute dominance of the segregating genes were made by two methods. That of Fisher, Immer, and Tedin [1932], amplified by Mather [1949], gave highly variable and statistically unreliable values. However, the results obtained indicate that the method developed by Comstock and Robinson [1952] was an efficient way to estimate genic dominance in these populations.

Two major loci were found to be largely responsible for locule number differences in these populations. It was possible to assign genotypic values to each of the nine major locule number genotypes. From these values, a theoretical mean absolute dominance estimate could be calculated, based on the two major genes only. This estimate, 0.66, was not markedly different from the 0.60 estimate obtained for all segregating genes by the method of Comstock and Robinson. Similarly, the potence estimate from the two major genes, 0.42, agrees closely with the estimate from the non-segregating generations, 0.43.

W. A. BECKER (State College of Washington, Western Washington Experiment Station, Puyallup, Washington).
459 **Computing Mortality in Experiments with Irrelevant Deaths.**

When computing mortality due to a specific cause, a, the experimentor has to make an adjustment for individuals that died of all other causes, b. Let us assume that a and b are independent and no individuals die of joint causes a and b. The mortality due to a can be computed by

calculating the proportion of individuals that did not die of a during each period of time, x, say the 24 hour periods in a yearly experiment. These proportions are multiplied and the product subtracted from one to give the desired mortality figure. A certain sample size has to be maintained during the experiment.

460 HANS ABPLANALP (University of California, Davis, California). **Effective Population Size Under Selection.**

Effects of artificial selection procedures on the relation between actual and effective population size are studied. The analysis is based on data from two generations of eleven closed chicken breeding flocks. Estimates of effective population size are based on methods derived by Crow and Morton (1954, Evolution 9: 202-214), contrasting expected sampling variance in gene frequency of an ideal population of N_e individuals with an estimate of the sampling variance derived from the data.

Four selection procedures are applied to the same breeding records, contrasting natural selection to 10 months of age, mass selection for a trait with high heritability (body weight), and two selection procedures for a trait with low heritability (egg production) where family averages are used to supplement information from individual production records.

461 R. W. ALLARD (University of California, Davis, California). **Effect of Genotypic—Environmental Interactions on the Prediction of Genetic Advance under Selection.**

Earliest and latest plants were selected in segregating generations of the wheat hybrid Baart × Ramona to determine the progress possible under different intensities of selection. In a parallel series of experiments, estimates were made of various genetic and environmental components of variance. Predictions of genetic advance under selection based on these estimates often did not correspond closely to the actual genetic advance. Although conspicuous interactions between genotype and environment occurred in these materials, these interactions did not appear to be the main cause of the failure of the predictions. The disturbing effect of a major gene on the selection differential, and nonallelic gene interactions, unaccounted for in the prediction equation, were advanced as more probable explanations for the discrepancy.

462 E. NOVITSKI AND E. DEMPSTER (Oak Ridge National Laboratory, Oak Ridge, Tennessee, and Department of Genetics, University of California, Berkeley, California). **Analysis by Automatic Computer of Data from Laboratory Populations of Drosophila Melanogaster.**

In some cases heterosis (hybrid vigor) is inferred because a population appears to contain more heterozygotes than can be accounted for by ordinary genetic theory. A mathematical model based on a maximum likelihood formulation has been set up that allows an estimate to be made of the degree to which irregularities in genetic data required the assumption of heterosis. The analysis based on such a model is not practicable by ordinary methods, and for this reason an automatic computer (the Oracle at Oak Ridge National Laboratory) was used. The output of the computer showed that the assumption that heterosis is responsible for the increased frequency of the heterozygotes is unwarranted, and that a reduced viability of homozygous classes will equally well explain such data. This result is of some interest because it shows that one of the most commonly applied tests for heterosis, a statistical test for an excess of heterozygotes in a population, is actually ambiguous.

463 L. H. KOOPMANS AND DOROTHY C. LOWRY (University of California, Berkeley, California). **Estimation of Mean and Variance of a Distribution of Binomial Probabilities.**

The following binomial probability situation is frequently encountered. Each individual in a population P has associated with it a probability p which reflects the individual's potential for inducing a characteristic C in each of a number of items it produces. Suppose that the distribution function of these probabilities is $F(p)$ with mean and variance π and σ_p^2. A random sample of k individuals is selected from P. For each i ($i = 1, 2, \cdots, k$), n_i items from the production of the individual are chosen at random and are tested for the presence of the characteristic C. On the basis of the resulting numbers of items possessing C we have derived unbiased estimates of π and σ_p^2. A natural estimate of π is

$$\hat{\pi} = \frac{1}{n} \sum_{i=1}^{k} X_i$$

where $n = \sum_{i=1}^{k} n_i$ and X_i is the observed number of items possessing

the characteristic C from the i^{th} individual. An unbiased estimate of σ_p^2, $\hat{\sigma}_p^2$, is

$$\hat{\sigma}_p^2 = \left\{\frac{n-k}{n-1}\left[\frac{n^2 - \sum_{i=1}^{k} n_i^2}{n}\right]\right\}^{-1}\left\{\sum_{i=1}^{k}\frac{X_i^2}{n_i} - \frac{k-1}{n-1}\left(\sum_{i=1}^{k} X_i\right) - \frac{n-k}{n(n-1)}\left(\sum_{i=1}^{k} X_i\right)^2\right\}$$

Exact expressions were derived for the variances of $\hat{\pi}$ and $\hat{\sigma}_p^2$.

464 ALBERT C. WALKER (International Minerals and Chemical Corporation, Woodland, California). **Balanced Incomplete Diallel Crosses.**

Progeny testing of genotypic values can be extended to a wider range of parents by use of an incomplete diallel. Comparison may be equalized and calculation simplified by application of the balanced incomplete block technique. A table of the number of males, number of females, and number of crosses per parent (complete to 200 progenies), necessary for balanced comparison, is presented. When the number of males is different from the number of females, only the smaller set gives complete balance. Cases where the larger set affords partial balance with only two associate classes are indicated.

Calculations of best estimates of parental values and their variances are detailed, with two examples involving glutamate content of progenies from sugar beet diallel crosses.

THE BIOMETRIC SOCIETY

Regional

E.N.A.R.

The Tenth Annual Meeting was held in Atlantic City, September 10–13, 1957. This was also a joint meeting with the American Statistical Association and the Institute of Mathematical Statistics.

Some of the session subjects and speakers were:

A. Methodology and Application of Response Surface Exploration: J. S. Hunter, R. M. DeBaun;
B. Forecasting Crop Yields: G. M. Kuznets, Y. Mundlak, H. F. Huddleston, R. Parr, L. Calvin;
C. Stochastic Processes in Biology: J. G. Huffman, E. Fix, E. K. Harris;
D. Genetic Statistics: W. R. Harvey, D. S. Robson, O. Kempthorne;
E. Fractional Replication: R. L. Anderson, J. D. Hromi, M. Carroll, C. Daniel;
F. Psychometric Statistics: H. Gulliksen, F. M. Lord;
G. Tolerance Limits: J. E. Walsh, E. A. Thomas, A. Lieberman;
H. Clinical and Field Trials in Public Health and Medicine: F. M. Hemphill, S. W. Greenhouse, J. Williams, P. Armitage;
I. Diet and Heart Disease: F. Moore, M. Robins, E. A. Lew.

Attendance at several sessions was 150 to 200. The annual business meeting was also held in Atlantic City. The minutes follow this report.

E.N.A.R. Minutes of Tenth Annual Business Meeting

The Tenth Annual Business Meeting of the Eastern North American Region of the Biometric Society was called to order at 2:00 p.m., September 12, in the Ambassador Hotel, Atlantic City. The presiding officer was Boyd Harshbarger, the Regional President. There were forty-four persons in attendance. The minutes of the Ninth Annual Business Meeting were read and approved.

S. M. Free, Chairman of the Regional Advisory Board, reported on a recommendation by that body that the ENAR By-laws be revised to make provision for a Regional President-elect. The reasons advanced by the RAB for the possible revisions were:

1. It would improve the continuity of the office of regional president. At present, a president is usually returned to office for a second term. At the end of that term a completely uninitiated person takes over the position.

2. The presidential honor (and duties) would be given to more people.

3. The president-elect could provide a needed over-all program chairman for ENAR's many meetings.

The amendments proposed by the RAB are as follows:

1. By-law 3. *Regional Committee*
   Change the first sentence to read:
   "The Regional Committee shall consist of the Regional President who shall serve as its chairman, the Regional President-elect, the Regional Secretary-Treasurer, and six ordinary members who shall serve three-year terms."
2. By-law 7. *Program Committees*
   Delete present By-law and substitute the following:
   "The Regional President shall appoint program committees to arrange meetings in connection with those of specific organizations or groups of organizations. The Regional President-elect shall act as program coordinator and shall establish appropriate liason with program committees of organizations with whom the Region meets. He shall report the tentative program plans for the coming year at the Annual Meeting."
3. By-law 8. *Election*
   Change the first sentence to read:
   "On or before August 1 each year the Regional Advisory Board shall submit nominees for Regional President-elect, Regional Secretary-Treasurer, and two nominees for each vacancy among ordinary members of the Regional Committee to the Regional Secretary-Treasurer."
4. By-law 8. *Election*
   Add the following sentence:
   "The Regional President-elect shall serve for one year and at the close of this term automatically become Regional President, subject to confirmation by the Council of the Society."

These amendments were voted on by those attending the meeting. The amendments passed without a dissenting vote.

It was also moved that the ENAR secretary be permitted to circulate ballots by mail for the election of a Regional President-elect for 1958. The RAB had already nominated for this position and it was

suggested that the ENAR secretary should send these ballots to the membership as soon as practicable after the present meeting.

A short report on the plans for each of the following meetings was given:

Gatlinburg meeting on April 10, 11, 12, 1958.
Indianapolis meeting after Christmas, 1957.
The Fourth International Biometric Conference and International Symposium on Biometrical Genetics to be held in Ottawa August 28 to September 2, 1958.

It was moved that the Chicago meeting with the ASA December 27–30, 1958, be designated the Eleventh Annual Meeting of ENAR.

Ralph Bradley was introduced as the new editor of *Biometrics*. He said that: the June issue would be out soon, the September issue on the Analysis of Covariance will be late, and the December issue will be out on time.

C. I. Bliss offered the following resolution to be suitably communicated to John W. Hopkins: Resolved that the Eastern North American Region of The Biometric Society express its sincere appreciation for the interest, enthusiasm, and leadership contributed by John W. Hopkins as editor of *Biometrics* and in his service to the Region in that capacity.

The resolution was unanimously subscribed to by the meeting.

R. G. Hoffman initiated a discussion concerning the problem of communication and interchange with the Medical Sciences. It was suggested that we should continue meeting with the FASEB and initiate further meetings with the medical clinical groups. It was suggested that perhaps a committee should be set up to implement such meetings. After considerable discussion of initiating further meetings of ENAR it was moved, seconded and passed that the Regional President should appoint a committee to explore and arrange with the advice and agreement of the Regional President-elect the feasibility of a joint meeting with the American Medical Association, the American College of Surgeons, or a similar group at the earliest effective and satisfactory date.

The meeting was adjourned at 3:05 p.m.

Respectfully submitted,
Arthur M. Dutton

Région Belge

Le 25 octobre 1957, la Société Adolphe Quetelet a eu le plaisir et l'honneur d'accueillir à sa tribune le Professeur Jerzy NEYMAN de l'Université de Californie de passage en Belgique ainsi que le Profes-

seur Robert CONSAEL de l'Université de Bruxelles. Le séance était organisée à l'Institut Agronomique de l'Etat à Gembloux et réunissait une nombreuse assistance de professeurs et d'étudiants ainsi que des membres de la Société, en présence de Monsieur HESPEL, Recteur de l'Institut. La conférence du Professeur NEYMAN portait sur un travail original qu'il a développé en collaboration de Mademoiselle SCOTT sur la Théorie probabiliste du développement des populations biologiques conçues comme conglomérations de familles. Ce brillant exposé donné en français a été suivi d'une discussion animée dirigée par le Président de la Société Adolphe Quetelet, Monsieur R. LAURENT. Le Professeur CONSAEL a ensuite parlé du Processus de Poisson à une ou deux variables aléatoires. Lors de la discussion qui suivit, le Professeur NEYMAN mit en lumière la parenté existant entre cet exposé et le sien, insistant sur la portée pratique de l'un et de l'autre.

Officers

E.N.A.R.

The following officers and committee members were elected by mail ballot by the members of the Eastern North American Region:—Regional President for 1958—Boyd Harshbarger; President-elect for 1958—Jerome Cornfield; Secretary--T. W. Horner; Regional Committee for 1958–1960 -D. B. Duncan and H. O. Hartley. The President-elect, who will become President in 1959, is a new officer. One of the duties of this officer will be to handle the meetings and programs of the *E.N.A.R.*

W.N.A.R.

The Regional officers for 1958 are as folows:—President—J. L. Hodges, Jr.; Secretary-Treasurer—Miss M. M. Sandomire; Regional Committee—1956–1958, Bernice Brown, L. Calvin; 1957 1959, W. J. Dixon, J. C. R. Li; 1958 -1960, D. G. Chapman and D. Wohlschlag.

NEWS AND ANNOUNCEMENTS

Members are invited to transmit to their National or Regional Secretary (if members at large, to the General Secretary) news of appointments, distinctions, or retirements and announcements of professional interest.

Marcus Kjeldsberg, University of Minnesota, who spent the last year as Special Research Fellow in the Epidemiological Research Unit of the Cardiovascular Department of Michael Reese Hospital in Chicago, has joined the staff of the Department of Tropical Medicine and Public Health, School of Medicine, Tulane University as Instructor in Biostatistics.

MEETING OF THE BIOMETRIC SOCIETY, E.N.A.R.

The Biometric Society (E.N.A.R.) will meet jointly with the Institute of Mathematical Statistics at Gatlinburg, Tennessee, Thursday, Friday, and Saturday, April 10, 11, and 12, 1958. Titles and abstracts, the latter in duplicate in the form published in Biometrics, for contributed papers for the Biometric Society should be sent to Donald A. Gardiner, Oak Ridge National Laboratory, Post Office Box Y, Oak Ridge, Tennessee, not later than March 10, 1958.

SYMPOSIUM ON BIOMETRICAL GENETICS AND THE FOURTH INTERNATIONAL BIOMETRICAL CONGRESS

The Symposium on Biometrical Genetics and the Fourth International Biometrical Congress will meet in Ottawa, Canada, from August 28 to September 2, 1958. May 1 is the deadline for pre-registration and for receiving papers and manuscripts. Queries regarding programmes and contributed papers are to be sent to Mr. M. J. R. Healy, Secretary, The Biometrics Society, Rothamsted Experimental Station, Harpenden Herts., England; queries concerning local arrangements should be sent to Mr. G. B. Oakland, Statistical Laboratory, Science Service Building, Central Experimental Farm, Ottawa, Canada.

NATIONAL SCIENCE FOUNDATION TRAVEL AID

The National Science Foundation will award individual grants to defray partial travel expenses for a limited number of American scientists

participating in the Special Session of the International Statistical Institute to be held in Brussels, Belgium, September 1–6, 1958.

Application blanks may be obtained from the National Science Foundation, Washington 25, D.C. Completed Application Forms must be submitted by March 1, 1958.

NEW COOPERATIVE PROGRAM IN BIOSTATISTICS

The Virginia Polytechnic Institute, Blacksburg, Virginia, and the Medical College of Virginia, Richmond, Virginia, have initiated a cooperative graduate training program in biostatistics through a grant from the National Institutes of Health. This program will be supported by the combined facilities of the two institutions and is designed to meet in part the critical demand for statisticians in the medical and public health professions. It is proposed to train both M.S. and Ph.D. candidates in statistics as well as to give post doctorate training in the field of statistics to holders of Ph.D. degrees in related fields.

Most of the course work will be given at the Virginia Polytechnic Institute, but certain specific courses relating to medical and public health statistics will be taken by the candidates at the Medical College of Virginia in Richmond. It is expected that in general a doctoral candidate will spend one year of his graduate study at the Medical College of Virginia.

To carry on this program, a staff of nine in the Department of Statistics at the Virginia Polytechnic Institute is being increased by one professor specifically trained in biostatistics, and the Medical College is adding a professor of biostatistics who will supervise the work in statistics at that division. The number of fellowships for candidates for graduate degrees now existing at the Virginia Polytechnic Institute is being increased in number as well as in amount to take care of this new field of study in biostatistics. In addition to these fellowships post doctorate fellowships are available under this new program for students in biology, medicine, and other related fields who wish to study statistics.

CONTENTS OF VOLUME 13